Second Language Prosody

van Maastricht, Lieke

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Link to publication in Tilburg University Research Portal

Citation for published version (APA):
van Maastricht, L. (2018). *Second Language Prosody: Intonation and rhythm in production and perception*. [Doctoral Thesis, Tilburg University]. [s.n.].

Lieke van Maastricht

Second Language Prosody

Intonation and rhythm in production and perception

Second Language Prosody

Intonation and Rhythm in Production and Perception

Lieke van Maastricht

Second Language Prosody
Intonation and Rhythm in Production and Perception

Lieke van Maastricht
PhD thesis
Tilburg University, 2018

TiCC PhD series No. 59

ISBN: 978-94-6295-903-3
Print: Proefschriftmaken
Cover design: Hans Westerbeek

Second Language Prosody

Intonation and Rhythm in Production and Perception

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan Tilburg University
op gezag van de rector magnificus,
prof. dr. E.H.L. Aarts,
in het openbaar te verdedigen ten overstaan van
een door het college voor promoties aangewezen commissie
in de aula van de Universiteit
op woensdag 9 mei 2018 om 16.00 uur
door

Lieke Joanne van Maastricht

geboren op 30 juni 1986 te Utrecht

Promotores

Prof. Dr. Emiel Krahmer
Prof. Dr. Marc Swerts

Overige commissieleden

Prof. Dr. Ad Backus
Prof. Dr. Aoju Chen
Dr. Núria Esteve-Gibert
Dr. Sónia Frota
Prof. Dr. Ineke Mennen

Financial support: Tilburg University, Prins Bernhard Cultuurfonds (grant 40005750/HEV/ILE) and the Jo Kolk Studiefonds.

Contents

1

1

General Introduction

When hearing another person speak from a distance, or maybe from another room, we often cannot hear exactly what is being said. However, we may be able to perceive aspects of the overall melody and tempo of the speech, and, interestingly, these so-called prosodic characteristics of speech seem to contain sufficient information to enable listeners to tell which language is being spoken, such as for example Dutch and Spanish (Ramus, Dupoux & Mehler, 2003; Ramus & Mehler, 1999). Apparently, we intuitively know what the melodic and rhythmic properties of different languages are, and are capable of discriminating between them, basing ourselves on these characteristics. And this does not only hold for adult listeners; even babies as young as five days old are able to discriminate between languages based on these prosodic features (Nazzi, Bertoncini & Mehler, 1998). It seems like the melodic and rhythmic characteristics of a language play a substantial role in what this language overall sounds like. Yet, how foreign language (L2) learners acquire these essential aspects of a language has remained largely unexplored and is the topic of this dissertation.

It is interesting to observe that the fact that languages differ in their prosodic properties receives very little or no attention in L2 classrooms. While pronunciation exercises are a common feature of formal L2 education, they are often aimed at practicing the production of consonants and vowels, in other words, the phonemes of a language, and not the production of higher-level properties, such as melody or tempo. Educational L2 methods tend to focus on accurate pronunciation and highlight those individual sounds that are difficult to produce for a certain group of mother tongue (L1) speakers. For instance, textbooks aimed at Dutch learners of Spanish generally emphasize the phoneme /θ/, as used in *cinco* (/θinko/, 'five'), because this segment is not part of the Dutch phoneme inventory. Generally lacking though in educational materials, is any information about those L2 sounds that supersede the segmental level: the suprasegments, or prosody, of a language, which include its rhythm and overall melody, also called intonation. Most handbooks for L2 Spanish mention word stress and explain how to correctly apply prominence at the lexical level based on the rules that exist for this purpose, because stress can be contrastive at the word level in Spanish (*preSENto* does not mean the same thing as *presenTÓ*), but to the best of our knowledge, there are no general L2 learning methods that specifically address Spanish sentence intonation and/or rhythm and explain to L2 learners how to acquire these language-specific characteristics. Unfortunately, this appears to hold for L2 manuals in general.

Perhaps the absence of prosodic features in L2 educational material can be explained by the fact that research on L2 acquisition, until relatively recently, was dedicated foremost to the learning of syntax, semantics and segments, which might be due to the (implicit) view that these linguistic areas are more essential for communication. Moreover, the majority of the studies that were dedicated to L2 prosody production, examined suprasegmental features as a part of pronunciation, and therefore only described what kinds of errors L2 learners tend to commit. For instance, Ramírez Verdugo (2005) showed that Spanish learners of English often use intonation patterns that are typical of their L1 but not of their L2 when they use intonation to express certainty. The fact that the role of prosody remained relatively under-represented in L2 acquisition studies is also reflected by the number of studies on the perception of L2 prosody by L1 listeners. However, it is interesting to investigate to which extent prosodic cues contribute to listeners' perceptions of how non-native L2 speakers sound, or how difficult it is to understand them. The relative contribution of different prosodic cues to these different types of L1 perceptions is a topic that is especially deserving of further exploration.

In the present dissertation, we therefore study the production and perception of L2 intonation and rhythm by Dutch leaners of Spanish and Spanish learners of Dutch. In contrast to most prior work, we do not only examine how the L2 prosody of these two learner groups differs from the L1 norm, but also try to explain this deviance by including additional factors in our designs, for example, proficiency level, learning direction, transfer direction, and syllable complexity. In addition, we approach L2 prosody production from a functional perspective; while the production of target prosodic features should be physiologically attainable for most L2 learners, the application of these features in the correct context is another matter altogether. In our perception studies, we explore whether errors in prosody production by L2 learners do not only affect L1 listeners' judgments about the L2 speaker, but also whether they can lead to actual miscommunication. In addition, we investigate which prosodic properties affect different kinds of L1 perceptions most, for instance, which cue contributes most to perceptions of non-nativeness: intonation or rhythm? In doing so, we aim to bridge the knowledge gap that exists about L2 prosody acquisition, to further L2 teaching, and to facilitate communication between L1 and L2 speakers. Before turning to the design of the current dissertation, we give a more detailed explanation of some of the concepts that are theoretically relevant in this context.

1.1 Prosody

Rietveld and Van Heuven (2009, p. 277) define prosody as "all sound properties of an utterance that are not related to those of its vowels and consonants". They explain that the ancient Greek, who coined the word *prosody*, thought that the segments conveyed most of the actual content of the utterance, whereas the prosody, or suprasegments, conveyed its intonation and rhythm, which were considered less relevant to the meaning a speaker wanted to convey. While the idea that the phonemes contain the essence of the message seems reasonable, it is not true that prosody cannot convey meaning at all (Ladd, 2008). Some prosodic phenomena indeed do not express any meaning in and of themselves; for instance, the decline of our pitch in the course of an utterance is the physiological result of the fact that the pressure on our vocal cords decreases as we exhale (Gussenhoven, 2004). Similarly, prosodic cues can play a role in the conveyance of speaker identity (Mennen, 2007), for example, the speech of L1 speakers of Southern Californian is characterized by a high rising tone, sometimes also referred to as 'uptalk' (Barry, 2007). However, sometimes prosody does convey (part of) the meaning of the message: In tone languages such as Mandarin, prosody is contrastive at the word level: it serves to distinguish lexical meaning in words that are produced with exactly the same consonants and vowels. For instance, the Mandarin word *ma* means 'mother' when pronounced with a level high tone, but 'horse' when using a tone that starts out high, then dips and then rises again. An example of how prosody can change the meaning of a sentence that might be more intuitive for L1 speakers of non-tonal languages is sarcasm; a discourse function that can be expressed by means of intonation and where one in fact means the opposite of what the utterance would normally convey (González Fuente, 2017). This is exemplified in (1a), where Lydia's comment is generally characterized by a falling intonation. When used in a non-sarcastic manner, as in example (1b), this utterance prototypically is produced with a rising tone.

(1a) John and Lydia are out to dinner. John accidentally spills some soup on his shirt.
Lydia: "How graceful you are!"

(1b) John and Lydia are taking a dancing class. John shows some real talent.
Lydia: "How graceful you are!"

Likewise, a change in intonation pattern may transform a declarative sentence into an interrogative one: In Spanish, a language in which word order changes are not a prerequisite to mark the difference between questions and declaratives as is the case in Dutch and English, an utterance such as *Habla español* ('He/She speaks Spanish') produced with a falling intonation pattern tends to be interpreted as the former, while a rising intonation would indicate the latter ('Does he/she speak Spanish?').

Another frequently used function of prosody in languages such as English and Dutch, and one that is further explored in several chapters of this dissertation, is the highlighting of new or contrastive information in an utterance, also called focus marking, by means of intonation. As shown in example (2), the utterance *JAN woont in Amsterdam* (small capitals indicate emphasis) seems a perfectly adequate response to utterance (2a), but is less appropriate as an answer to utterance (2b). Conversely, (2d) would be a good response to (2b), but should not be used as an answer to (2a).

(2a) Peter woont in Amsterdam, toch? ('Peter lives in Amsterdam, right?')

(2b) Waar woont Jan? ('Where does John live?')

(2c) JAN woont in Amsterdam. ('JOHN lives in Amsterdam.')

(2d) Jan woont in AMSTERDAM. ('John lives in AMSTERDAM.')

In order to investigate prosodic constructs such as intonation and rhythm, the following four properties of speech are generally measured:

⇒ *intensity* (the volume, or loudness, of speech)

⇒ *pitch* (corresponding to its fundamental frequency or F0, and the low or highness of tones)

- ⇒ *length* (the duration of a certain segment, be it syllable, intonational phrase, or utterance)
- ⇒ *timbre* (the quality of the sound, i.e., the spectral characteristics of speech

Figures 1 and 2 visualize the prosodic analysis of these four parameters in Praat (version 6.0.20, Boersma & Weenink, 2016), which is the program used for the scientific analysis of speech in this dissertation. **Figure 1** does so for the utterance *JAN woont in Amsterdam* ('JOHN lives in Amsterdam'), while **Figure 2** depicts the analysis of the utterance *Jan woont in AMSTERDAM* ('John lives in AMSTERDAM'). The top layer of both figures shows the waveform, or oscillogram. On the right y-axis, the intensity of the speech signal is measured in decibel, visualized by means of the blue line. On the left y-axis, the fundamental frequency of speech is measured in hertz, depicted in the figure by the green line. The duration of the speech signal is measured on the x-axis, in seconds. The spectral characteristics of the speech signal are visualized in the middle part of the figure, behind the lines representing intensity and pitch. The bottom part of the figures depicts the three layers, or tiers, that have been used for coding or annotating the speech signal: the transcription in tier 1 is at the word level, the one in tier 2 at the syllable level, and the one in tier 3 at the phoneme, or segment, level.

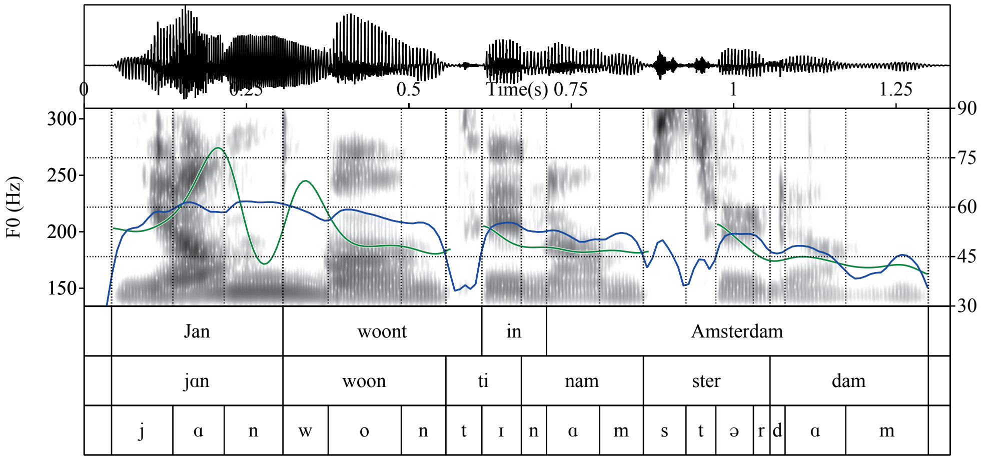


Figure 1. Waveform, spectrogram, F0 and intensity contour, and annotations for the Dutch utterance *JAN woont in Amsterdam,* 'JOHN lives in Amsterdam'.

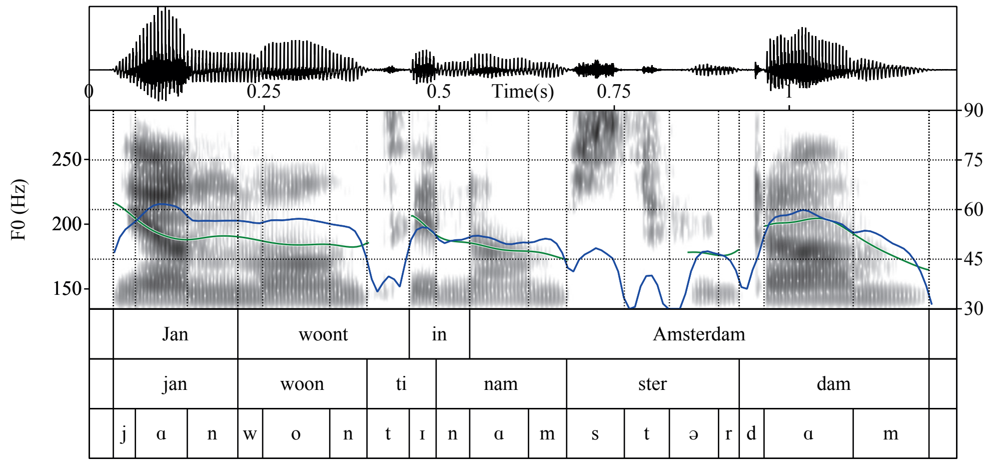


Figure 2. Waveform, spectrogram, F0 and intensity contour, and annotations for the Dutch utterance *Jan woont in AMSTERDAM,* 'John lives in AMSTERDAM'.

Adding to the complexity of measuring and defining constructs such as intonation or rhythm is the fact that their production is characterized by interplay between all four prosodic cues mentioned above. For instance, the concept of using intonation to make certain parts of the utterance more prominent, as explained in examples (2c) and (2d) in the context of focus marking, is a complex construct: depending on the language, all four of the prosodic cues may be used to convey phrasal emphasis, as is the case in Dutch (Sluijter, 1995, Chapter 3). Interestingly, languages do not only differ in the use of cues that they employ to realize constructs such as prominence, they also differ in the rules that govern its use. For example, in Dutch any word in the utterance may be prosodically emphasized to make it more prominent and elements may be deaccented to mark their givenness (Krahmer & Swerts, 2001; Nooteboom & Kruyt, 1987), while in Spanish almost all content words tend to be accented, with the last content word of the utterance receiving nuclear stress (Face, 2001; Hualde, 2005; Zubizarreta, 1998). As a result, word order changes are often employed in Spanish to ensure that the new or important part of the utterance is mentioned last and thereby emphasized (Hoot, 2012), while in Dutch, which has relatively fixed word order, emphasis is placed there where it is necessary to make a certain element of the utterance stand out.

Given that Dutch and Spanish differ in this respect, it is interesting to examine how L2 learners of these languages who have the other language as the L1 acquire the necessary prosodic cues and learn to apply them in the right communicative context. In the following section, we will briefly review the literature on the production and perception of L2 prosody, before detailing the contents and design of this dissertation.

1.2 Production and perception of L2 prosody

Previous studies on L2 prosody production mostly focused on descriptions of specific errors made by L2 learners. In general, these studies reported that the L1 of the language learner tends to influence the L2 and that learners tend to copy prosodic features from their L1 to their L2 (e.g., Rasier, 2006). This process, which is commonly referred to as '(linguistic) transfer' (Ellis, 1994), has mostly been investigated focusing on more formal aspects of prosody that in themselves do not convey a specific meaning, such as pitch range. For instance, Willems (1982) compared Dutch and English and found that L1 speakers of British English tend to employ a larger pitch range than Dutch learners of English. Relatively few studies investigated transfer effects with respect to prosodic features that have a communicative function, such as the marking of utterance boundaries by means of intonation (e.g., Swerts & Zerbian, 2010). Unfortunately, few of these studies included proficiency level as a factor in their design. From a didactical perspective, it would be interesting to investigate whether the acquisition of the prosodic properties of the L2 takes place from the beginning of the L2 acquisition trajectory or not, as well as to determine whether L2 learners are capable of acquiring the target prosody at all or whether the influence from their L1 proves too strong. This kind of information also has relevance for L2 didactical methods, which may use it to ensure that the right skills are offered to students at the right moment in their L2 education.

Finding out more about the type of prosodic errors that L2 learners make is important, because previous research has shown that we are very sensitive to the suprasegmental properties of speech, not only in the L1 (e.g., Cutler, 1976, Terken & Nooteboom, 1987), but also with respect to non-native speakers. L1 listeners usually are very proficient at discriminating between L1 and L2 speakers based on their pronunciation. Even when all segmental cues are removed, L1 listeners are capable of distinguishing between L1 and L2 speakers, using only the prosodic cues available to them (e.g., Van Els & De Bot, 1987). Furthermore, it has been shown that when L2 learners produce atypical prosodic patterns, this affects L1 listeners' perceptions of how foreign an L2 speaker sounds, how difficult L1 listeners find it to understand the message, and how many errors they make in actually processing it (e.g., Munro & Derwing, 1999). Interestingly, not all of these aspects of perception are influenced to the same extent and it seems reasonable to assume that there might be a hierarchical structure concerning the severity of different types of prosodic errors (e.g., an error in prominence marking, such as exemplified in (2) vs. speaking with more intensity or using a larger pitch range than normal for that language), as well as errors in different prosodic cues (e.g., atypical pitch production vs. atypical vowel quality). Therefore, one of the aims of the present dissertation is to investigate which type of prosodic errors contributes most to which types of perceptions.

In sum, while L1 speakers tend to have clear intuitions about both the production and perception of L2 prosody (Mennen, 2007), many questions remain to be answered about the factors that underlie the acquisition process itself, as well as the effect of prosodic errors in L2 speech on communication between L1 and L2 speakers. Hence, the present dissertation aims to study these aspects of L2 prosody production and perception.

1.3 The current dissertation

The studies reported in this dissertation all focus on L1 and/or L2 Spanish and Dutch and are centered around the following four main research questions:

RQ1 Do the differences between Dutch and Spanish in the way they use intonation to mark focus lead to transfer effects in the L2 speech of learners of these languages? If so, does proficiency level influence this effect?

RQ2 How does deviance in the production of intonation to mark focus by L2 learners influence L1 listeners' perceptions of their speech concerning its non-nativeness, the ease with which it is understood, and actual processing?

RQ3 Does the direction in which L2 acquisition takes place affect the successful attainment of speech rhythm by Spanish learners of Dutch and Dutch learners of Spanish?

RQ4 What is the relative contribution of intonation, rhythm and speech rate in speech produced by L2 learners on L1 listeners' perceptions of non-nativeness and ease of understanding?

In order to answer these questions, the current dissertation is set up using a double diptychal structure: **Chapters 2 and 3** are dedicated to *intonation*, while **Chapters 4 and 5** contain studies that are centered on *rhythm*. Within both diptychs, the first chapter always concerns the *production* of L2 prosody, whereas the second explores the *perception* of prosodic deviance in the L2 by L1 listeners.

Thus, in **Chapter 2** we report the results of an experiment in which Dutch learners of Spanish and Spanish learners of Dutch produced utterances that were varied in focus structure in order to investigate whether both learner groups were able to produce native-like intonation in varying contexts and whether the extent to which they were successful was affected by their overall L2 proficiency. Another aim of the study was to determine whether the intonational structure of the L2 of the participants could also influence the intonation of their L1, in other words, whether there is more than just one transfer direction. **Chapter 3** explores how deviance in the use of intonation to mark focus by L1 Dutch and Spanish learners of Dutch affects four different measures of L1 Dutch perceptions.

The study reported in **Chapter 4** examined whether Dutch learners of Spanish and Spanish learners of Dutch were equally capable of acquiring native-like speech rhythm, with the aim of determining whether learning direction, as well as syllable structure complexity, affects this process. This study is an almost identical replication of Prieto, Vanrell, Astruc, Payne & Post (2012), and bases its prediction that the speech rhythm of Dutch is more difficult to acquire for Spanish learners than vice versa on Eckmans' Markedness Differential Hypothesis (1977; 2008). In **Chapter 5** both prosodic cues investigated in the preceding chapters, that is, rhythm and intonation, as well as most of the previously used the perception measures, are combined in one design. More specifically, the relative contribution of deviance in speech rate, intonation, and rhythm in speech by Spanish learners of Dutch to L1 Dutch perceptions of accentedness and comprehensibility is addressed. In contrast to the study reported in

Chapter 3, in which speech stimuli were created using the original utterances of the speakers, without manipulating the signal itself, in Chapter 5 speech transplantation and resynthesis techniques were used to create the speech stimuli. Due to this method, its design resembles the combination of errors that are typically made by L2 learners. The final chapter of this dissertation contains the answers to the four main research questions stated above, as well as a general discussion of the dissertation, including its theoretical and practical implications and suggestions for future research.

As a final note, we would like to state that the chapters of this dissertation consist of self-contained studies, two of which have been published in, and two of which are submitted to peer-reviewed scientific journals. As such, they all include their own abstract, introduction, discussion, and reference list, which entails that some overlap between respective chapters is unavoidable. Since the different studies have all been published in or submitted to different journals, small variations in style, as well as statistical procedures, can be expected.

References

Barry, A. (2007). *The form, function, and distribution of high rising intonation in Southern Californian and Southern British English* (Doctoral dissertation). University of Sheffield, Sheffield, United Kingdom.

Boersma, P., & Weenink, D. (2016). Praat: Doing phonetics by computer (Version 6.0.20) [Computer software]. Retrieved from http://www.fon.hum.uva.nl/praat/

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Attention, Perception, & Psychophysics*, *20*(1), 55-60.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning 27*(2), 315–330.

Ellis, R. (1994). *The study of second language acquisition.* Oxford: Oxford University Press.

Face, T. L. (2000). Prosodic manifestations of focus in Spanish. *Southwest Journal of Linguistics*, *19*, 45–62.

González Fuente, S. (2017). *Audiovisual prosody and verbal irony* (Doctoral dissertation). Universitat Pompeu Fabra, Barcelona, Spain.

Gussenhoven, C. (2004). *The phonology of tone and intonation.* Cambridge, United Kingdom: Cambridge University Press.

Hoot, B. (2012). *Presentational focus in heritage and monolingual Spanish* (Doctoral dissertation). University of Illinois at Chicago, Chicago, United States.

Hualde, J. I. (2005). *The sounds of Spanish.* Cambridge, United Kingdom: Cambridge University Press.

Krahmer, E., & Swerts, M. (2001). On the alleged existence of contrastive accents. *Speech Communication*, *34*, 391–405.

Ladd, D. R. (2008). *Intonational phonology.* Cambridge, United Kingdom: Cambridge University Press.

Mennen, I. (2007). Phonological and phonetic influences in non-native intonation. In J. Trouvain, & U. Gut (Eds.), *Non-native prosody: Phonetic description and teaching practice* (pp. 53-76). Berlin, Germany: Mouton de Gruyter.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *49*(s1), 285-310.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human perception and performance*, *24*(3), 756-766.

Nooteboom, S. G., & Kruyt, J. G. (1987). Accents, focus distribution, and the perceived distribution of given and new information: An experiment. *Journal of the Acoustical Society of America*, *82*, 1512–1524.

Prieto, P., Vanrell, M., Astruc, L., Payne, E., & Post, B. (2012). Phonotactic and phrasal properties of speech rhythm: Evidence from Catalan, English, and Spanish. *Speech Communication*, *54*, 681–702.

Ramírez Verdugo, M. D. (2005). The nature and patterning of native and non-native intonation in the expression of certainty and uncertainty: Pragmatic effects. *Journal of Pragmatics*, *37*(12), 2086-2115.

Ramus, F., Dupoux, E., & Mehler, J. (2003). The psychological reality of rhythm classes: Perceptual studies. In M. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 337–342). Barcelona: Universitat Autonomá de Barcelona.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *The Journal of the Acoustical Society of America*, *105*(1), 512-521.

Rasier, L. (2006). *Prosodie en vreemdetaalverwerving. Accentdistributie in het Frans en het Nederlands als vreemde taal* (‘Prosody and foreign language learning. A distributional study of pitch accent in Dutch and French as a foreign language’, in Dutch) (Doctoral dissertation). Université catholique de Louvain, Louvain, Belgium.

Rietveld, T., & van Heuven, V. J. (2009). *Algemene fonetiek* (‘General phonetics’, in Dutch). Bussum, The Netherlands: Coutinho.

Sluijter, A. (1995). *Phonetic correlates of stress and accent* (Doctoral dissertation). Leiden University, Leiden, The Netherlands.

Swerts, M., & Zerbian, S. (2010). Intonational differences between L1 and L2 English in South Africa. *Phonetica*, *67*, 127–146.

Terken, J., & Nooteboom, S. G. (1987). Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes*, *2*, 145–163.

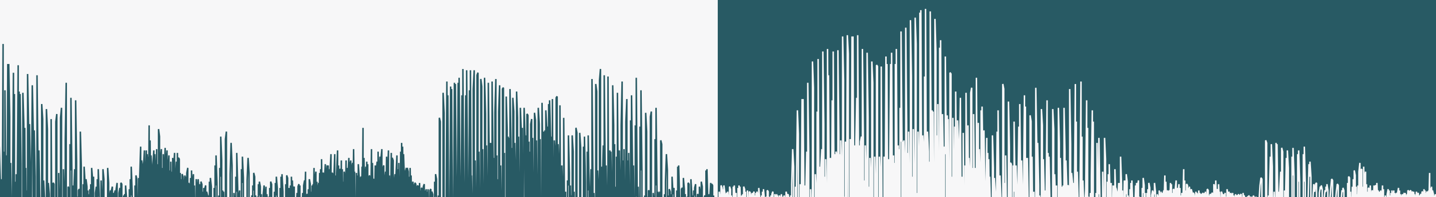
Van Els, T., & De Bot, K. (1987). The role of intonation in foreign accent. *The Modern Language Journal*, *71*(2), 147-155.

Willems, N. (1982). *English intonation from a Dutch point of view*. Dordrecht, The Netherlands: Foris.

Zubizarreta, M. L. (1998). *Prosody, Focus, and Word Order*. Cambridge, Maryland, United States: MIT Press.

2

intonation production



2

Prominence patterns in a second language:

Intonational transfer from Dutch to Spanish and vice versa*

Abstract

This research describes the production of prosodic cues to mark information structure in Spanish and Dutch. It compares speech by native (L1) and second language (L2) speakers and investigates prosodic transfer from the L1 to the L2, L2 proficiency as a factor in transfer effects, and transfer from the L2 to the L1. The results confirm that Spanish and Dutch natives use different prosodic cues to mark information status. Comparison of L1 and L2 data reveals that these prosodic differences lead to transfer from the L1 to the L2. The proficiency level of the speaker modulates transfer effects. To some degree, pitch accents used to mark focus appear to be transferred from the L2 to the L1 as well.

* **This chapter is based on:** Van Maastricht, L., Krahmer, E., & Swerts, M. (2016). Prominence patterns in a second language: Intonational transfer from Dutch to Spanish and vice versa. *Language Learning, 66*(1), 124-158.

2.1 Introduction

Most second language (L2) learners know that in order to come across as a native speaker they must attend to a whole range of linguistic features. Linguistic difficulties of L2 learners clearly manifest themselves in the articulation of specific sounds: Although a Spanish sentence like *Beatriz Gómez juega al voleibol* ("Beatriz Gómez plays volleyball") is grammatically accurate, a native speaker of Dutch may have difficulties pronouncing it correctly because certain orthographic representations correspond to different phonetic segments in Dutch and Spanish. For example, in Dutch, the <g> is generally pronounced as [x] instead of [ɡ] (or [ɣ] in intervocalic position), as is correct in this context, and the <v> and <j> are not pronounced as [b] and [x], but as [v] and [ʎ]. The sound [θ], as used for the <z> of *Beatriz* when pronounced by a speaker of Castilian Spanish (the variety of Spanish under investigation in the current study), does not exist in Dutch and is therefore often mispronounced as [s] or [z]. When languages have diverging phonological systems, as evident for Dutch and Spanish, this often results in a non-target pronunciation in the L2. This is caused by the fact that learners copy certain, in this case phonological, features of their native language (L1) onto the language they are acquiring. This process is commonly referred to as (linguistic) transfer (Ellis, 1994, p. 28).

Previous research has shown that transfer from the L1 to the L2 can be related to syntactic (Robertson, 2000), semantic (Jarvis, 2000; Jiang, 2004), or segmental (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Flege, Schirru, & MacKay, 2003; Guion, 2003) components of language. However, few studies have investigated another aspect that may form a possible distinguishing factor between L1 and L2 speakers: prosodic (or suprasegmental) features, such as intonation, stress, and rhythm. These are often considered to be language features that are difficult to manipulate consciously. Therefore, the acquisition of these prosodic features generally receives little attention in educational programs (in contrast to the acquisition of individual segments, which is usually discussed quite extensively). Suprasegmentals also belong to a class of features that tend to be minimally represented in writing systems, which could be another reason why they receive less attention in (textbook) approaches to L2 acquisition. Of course, some aspects of prosody do not require learning because they emerge automatically as a result of articulatory constraints on speaking. For instance, the declination effect that pitch tends to lower in the course of an utterance is often seen as a natural consequence of a decreasing air pressure; similarly, some pauses are inserted in speech to give speakers the opportunity to draw breath. However, languages do differ in the way they use prosody to encode specific, often communicative, functions.

Consequently, it would seem important that this aspect of a L2 should be investigated and taken into account during the acquisition process, yet this is seldom the case. This is unfortunate, because the importance of intonation in spoken communication cannot be denied (Caspers & Horloza, 2012; Hahn, 2004; Ladd, 2008); it plays a key role in the regulation of discourse, the marking of new and given information, the conveyance of speaker identity (Mennen, 2007), and the perceived quality of the speech (Swerts & Marsi, 2012). Therefore, the use of an inappropriate intonation pattern may lead to miscommunication or incomprehensibility (Mennen, 2007; Munro & Derwing, 1999; Terken & Nooteboom, 1987). Because the goal of most L2 learners is to successfully communicate in a foreign language, more research on possible transfer effects in their use of intonation is highly relevant.

The present research focuses on the existence of possible prosodic transfer effects and takes a functional approach by concentrating on a fundamental function of intonation (Cutler & Ladd, 1983; Swerts & Zerbian, 2010): the use of pitch accents in the marking of focus. This prosodic feature is investigated in speech produced by L1 and L2 speakers to determine whether transfer occurs. Additionally, it is examined to what extent intonational transfer is modulated by proficiency level. Furthermore, the direction in which possible transfer effects might take place is investigated: Can the L1 only influence the L2 or is a reversed effect also possible?

2.2 Theoretical background

2.2.1 Functional intonation: Marking prominence

Generally, a distinction is made between intonation languages and word-order languages or, as Vallduví (1991) referred to them, plastic and non-plastic languages. In the former, intonation is used to mark the information status of a certain element by means of pitch accents (rises in the melodic patterns that make words sound more prominent), while in the latter this is reflected by its position in the sentence and the fact that the nuclear accent is generally placed at that, usually fixed, position. Dutch is taken to be a plastic language (Rasier, 2006). Roughly, this entails that new and contrastive information tend to be accented, while given information is usually deaccented (Krahmer & Swerts, 2001; Nooteboom & Kruyt, 1987; Terken, 1984). Deaccentuation occurs when "a word that we might expect to be accented fails to be accented in a context where it has recently been used or where the entity to which it refers has recently been mentioned" (Ladd, 2008, p. 175). For instance, in the subordinate clause in example (1), "Mary" is deaccented in favor of "like" (SMALL CAPITALS represent accentuation).

(1) John came to the party with MARY, even though he knows I don't LIKE Mary.

Spanish, however, is a non-plastic language, which is more constrained in its use of accent distribution to reflect the information status of the sentence elements. In Spanish, the pitch pattern of utterances in broad and narrow focus generally is identical (Face, 2002) and the nuclear accent is placed on the last content word of the intonational phrase (Hualde, 2005; Zubizarreta, 1998). Deaccentuation does not commonly occur in Spanish (Cruttenden, 1993), as content words are usually accented (Hualde, 2009). When a non-final element of the sentence receives narrow focus, often a word-order modification (among other strategies, see Hoot, 2012) ensures that the nuclear accent falls on the focused element, as shown in example (2a), adapted from Face (2000, p. 46). In example (2a), the direct object "boek" is most prominent. In order for the direct object to receive the nuclear accent in Spanish, it is moved from its canonical position directly behind the verb to the right periphery of the utterance (cf. 2b). This positioning is not necessary in Dutch, where nuclear stress can be placed anywhere in the sentence, depending on the context. Thus, in Dutch pitch accent distribution appears to be driven by pragmatics, whereas it is driven by word status in Spanish.

(2a) *Juan gaf het* BOEK *aan María.*
John gave the BOOK to Mary.
'John gave the BOOK to Mary'.

(2b) *Juan le dio a María el* LIBRO.
John to+her gave to Mary the BOOK.
'John gave the BOOK to Mary'.

In light of this difference, it is expected that a comparison of speech by L1 speakers of Dutch and Spanish in contexts with a fixed word order (e.g., a noun phrase [NP]), but with varying information status types, shows different intonation patterns for the two groups. In Spanish, the canonical word order within the NP (i.e., [determiner] + noun + adjective) can usually not be varied, with the exception of those cases in which a change in word order is possible, but also entails a change of meaning (e.g., *un coche nuevo* means "a new car," while *un nuevo coche* means that the car might be old, but that it is new to me), so in Spanish NPs both the noun and adjective are expected to be accented, with the main pitch accent at the end of the intonational group, irrespective of the information status of the pre-final element (e.g., *guante* ROJO, "red glove"). In Dutch, however, prominence depends on which element is new or contrasting in the context: for example, RODE *wanten* ("red mittens") if the pre-final element is new or contrasting with earlier referents and *rode* WANTEN if the final element is new or contrasting with earlier referents. Because Dutch and Spanish differ in this way, they are especially suited for a cross-linguistic analysis, because a difference between these two languages might manifest itself in transfer from the L1 to the L2.

2.2.2 Prosodic transfer in functional contexts

In contrast to other types of transfer, prosodic transfer in a functional context remains under-investigated. Most studies that did look into prosodic transfer in a functional context show that the observations from studies on linguistic transfer in general can be generalized to prosodic transfer (e.g., Gut, Pillai, & Mohd Don, 2013; Gut & Pillai, 2014; Trofimovich & Baker, 2006; Ulbrich, 2013; Willems, 1982). Ramírez Verdugo (2002), for instance, provides a qualitative analysis of the intonation patterns used by Spanish learners of English in statements, answers, and different types of questions (on prosodic transfer in yes/no questions, see Gabriel & Kireva, 2014). She reports that participants are able to produce L2 intonation patterns in questions and answers, but that they do not mark contrast or givenness prosodically and incorrectly use features of their L1 intonation in these contexts. Although this qualitative study offers interesting observations, further analysis is needed at the quantitative level to reinforce them. Nava and Zubizarreta (2010) and Zubizarreta and Nava (2011) also show that intermediate and advanced Spanish learners of English transfer L1 pitch accent patterns to their L2 when marking broad and narrow focus. In their experiment, which was performed by a larger sample than in Ramírez Verdugo's study, participants read aloud question-and-answer dialogues.

Rasier and Hiligsmann (2009) typologically compared the use of (de)accentuation to mark focus in Dutch and French utterances produced by L1 and L2 speakers with a design similar to the one used by Swerts, Krahmer, and Avesani (2002), in which participants described geometrical figures and their colors to each other. The design elicited utterances with varying information status, with either the color or the shape representing contrasting information (e.g., a red triangle followed by a blue triangle or a blue square followed by a blue triangle). Rasier and Hiligsmann also report the existence of transfer effects from the L1 to the L2, but signal that more transfer occurs in the data of French than Dutch learners. In general, Dutch learners of French produce intonation patterns that are also used in L1 French. The only nonnative aspect of their data seems to be the fact that they at times deaccentuate given elements, which is atypical of L1 French, in which deaccentuation does not occur. The French learners of Dutch seem to be less successful at producing target intonation patterns, as they generally fail to deaccent given information and use the same L1 intonation pattern in all contexts, independent of their information structure, which is typical of French. Their results show that the extent to which transfer occurs might be related to typological differences between the two languages. Using the Markedness Differential Hypothesis by Eckman (1977), Rasier and Hiligsmann show that a language with a structural accentuation rule, like French, seems to be easier to acquire than a language with accentuation rules based on the pragmatic context, like Dutch (see also Flege, 1995).

Swerts and Zerbian (2010) concentrated on information status marking and the signaling of continuation or finality in a list in L1 English, L1 Zulu, and L2 English spoken by L1 Zulu speakers. Although previous studies obtained similar results as Swerts and Zerbian in other subareas of L2 acquisition (e.g., Flege & Hillenbrand, 1984; Foote, 2010; Imai, Walley, & Flege, 2005), their research is rare because they incorporated proficiency level as a factor into a study on prosodic transfer. Their results indicate that highly proficient Zulu learners of English (with a C2 level in receptive language skills, according to the Common European Framework of Reference [CEFR] for languages by the Council of Europe, 2001) produced intonation patterns that were similar to those produced by L1 speakers of English in all contexts. They used pitch accents to mark focus and produced adequate boundary tones, while less proficient speakers (with a B2 proficiency level in receptive language skills) produced target boundary tones to signal finality and continuation, but did not use intonation to mark focus, which corresponds to what is typical of their L1. The only other study known to us that reports modulating effects of proficiency level on prosodic transfer is Reichle and Birdsong (2014), who compare event-related potentials elicited among 12 L1 and 24 L2 speakers of French producing semi-spontaneous sentences in varying focus conditions. The 24 L2 learners were divided into two proficiency groups based on the average of their self-reported speaking, reading, listening, and pronunciation proficiency scores. Their findings suggest that a higher proficiency level is associated with a lesser degree of prosodic transfer.

Although many previous studies provide interesting and valuable insights into several aspects of the L1–L2 relationship that seem to be relevant to L2 acquisition, it is difficult to compare them systematically because they use different data collection methods, analyze data in different ways, and use different types of stimuli. For instance, several studies base their analysis on listener-dependent prominence judgments made by the authors and/or others, rather than on independent acoustic measurements (e.g., Rasier & Hiligsmann, 2009; Swerts & Zerbian, 2010). Others were carried out using a relatively small sample (e.g., Gut & Pillai, 2014; Henriksen, Geeslin, & Willis, 2010; Mennen, 2004) or did not take into account the proficiency of the L2 learners (e.g., Gut & Pillai, 2014; Mennen, 2004; Ramírez Verdugo, 2002; Rasier, 2006; Rasier & Hiligsmann, 2009; Swerts et al., 2002). Additionally, varying ways of eliciting speech are reported, ranging from tasks eliciting semi-spontaneous speech (e.g., Rasier & Hiligsmann, 2009; Swerts & Zerbian, 2010; Swerts et al., 2002) to oral reading tasks eliciting less spontaneous, and possibly less natural, speech (e.g., Gut et al., 2013; Gut & Pillai, 2014; Nava & Zubizarreta, 2010; Ramírez Verdugo, 2002). Several studies have shown that spontaneous and read-aloud speech has significantly different prosodic characteristics (e.g., De Ruiter, 2015; Howell & Kadi-Hanifi, 1991; Swerts, Strangert, & Heldner, 1996). Because many L2 learners likely intend to use their L2 in communication, our study is designed to elicit semi-spontaneous, natural speech by means of a naming task to closely approximate the prosodic features used in real-life communication. Finally, to our knowledge, ours is the first study to compare Dutch and Spanish, which are especially suited for a contrastive analysis because they differ with respect to the way they use pitch accent placement to express information status in the manner explained above.

2.2.3 Bidirectional transfer

Another aspect related to prosodic transfer is the notion of bidirectionality, which has been defined as "the two-way interaction between the two linguistic systems of a second language user, i.e., native language influence on the second language and second language influence on the native language" (Pavlenko & Jarvis, 2002, p. 192). For instance, people who have lived abroad for a long time often pronounce L1 phonemes in a way that is characteristic of their L2, which may result in a nonnative accent. Bidirectional transfer in adults has been studied most extensively in the fields of syntax, semantics, and segmental pronunciation, while bidirectional transfer at the suprasegmental level has been studied very little (for an overview, see Pavlenko, 2000).

To our knowledge, the only study on prosodic transfer from the L2 to the L1 is Mennen (2004), dealing with the timing of non-final and pre-nuclear rises by highly experienced Dutch speakers of Greek. Her data, which were examined by means of acoustic analyses within the autosegmental-metrical framework (Pierre-humbert, 1980), found evidence for bidirectional transfer: Four of the five participants produced inaccurate F0 (fundamental frequency) timing both in their L1 and in their L2. These are intriguing results, but they were obtained for a small number of participants, who spoke the L2 on a highly regular basis, and looking at non-communicative aspects of prosody, so that it is unclear whether these effects generalize. In an attempt to contribute to the claims made in Mennen (2004) and other prior work, the current experiment has a large sample size (N = 124, with 45 L1 speakers and 79 L2 learners) and it involves an aspect of language in which the adequate production of intonation can be crucial for communication.

2.3 The current study

As is clear from the literature reviewed here, the topics of prosodic transfer, the modulating function of L2 proficiency, and the possibility of bidirectional transfer remain under-investigated, in comparison to the work on transfer in other linguistic areas. The innovative aspect of the current study lies in the fact that it contrasts two languages that have never been compared directly in this context in a design that encompasses all the factors that have been deemed relevant in previous work but that until now have only been investigated separately, on a relatively small scale, in nonspontaneous speech, and/or by using subjective non-acoustic analyses. Additionally, the current design is similar enough to the one used in prior studies (e.g., Gut et al., 2013; Rasier & Hiligsmann, 2009; Swerts & Zerbian, 2010) to enable an overall comparison of results, but is novel in the sense that it uses objective measurement techniques in combination with a large participant sample and focuses on the use of prosody in a functional context. The current study is set up to address four research questions (RQs):

(1) How do native speakers of Dutch and Spanish differ in the way they use prosodic cues to mark focus?

Based on previous work (e.g., Face, 2002; Hualde, 2005; Nooteboom & Kruyt, 1987; Rasier, 2006; Krahmer & Swerts, 2001), it is predicted that Dutch natives use pitch accent distributions to mark focus, while Spanish natives do not. Specifically, it is expected that the Spanish natives produce similar pitch accent distributions in all focus conditions, in which the second word carries the main accent of the NP. Contrastingly, Dutch natives are predicted to produce pitch accent distributions that reflect the focal status of the words in the NP in that a contrasting word is accented and a given word is deaccented. This prominence difference between accented and deaccented words is expected to be especially noticeable in the two conditions in which one of the words of the NP is contrasting and the other is given (narrow focus). The remaining two focus conditions, that is, the one in which both words have contrastive focus and the one in which both words are given, are expected to follow the default pitch accent distribution pattern in which the second word receives the main stress, but in which the prominence difference between the two words is less noticeable.

(2) Do the differences between native Dutch and Spanish in the way they use pitch accent distributions to mark focus lead to transfer effects in the L2 speech of learners of these languages?

Based on studies on prosodic transfer (e.g., Gut et al., 2013; Gut & Pillai, 2014; Ramírez Verdugo, 2002; Rasier & Hiligsmann, 2009; Swerts & Zerbian, 2010; Zubizarreta & Nava, 2011), it is predicted that transfer takes place to some extent in both learner groups.

(3) Are these possible transfer effects modulated by the L2 proficiency level of the participants?

As Swerts and Zerbian (2010), Zubizarreta and Nava (2011), and Reichle and Birdsong (2014) found evidence of a modulating effect of proficiency level on the degree of transfer, it is probable that the current study reproduces these results. For instance, Swerts and Zerbian report that, while proficient learners successfully acquire L2 boundary tones and pitch accents without transfer effects, less proficient learners still transferred characteristics of their L1 to their L2 in their use of pitch accents.

(4) Do transfer effects also take place bidirectionally, that is, from the L2 to the L1?

Bidirectional prosodic transfer effects have only been demonstrated once (Mennen, 2004), on a small scale and in a nonfunctional context. It might be that bidirectional transfer does not occur in the use of prosody in a functional manner as it does in less functional contexts. Additionally, it might be that bidirectional transfer only occurs in speech of highly advanced speakers, as Mennen's participants used the L2 on a daily basis and had at least 12 years of L2 experience. Therefore, it remains improbable that our learners, who are less proficient in comparison to those of Mennen (2004), transfer prosodic characteristics from the L2 to the L1.

2.4 Method

2.4.1 Participants

Throughout the study, four groups of participants are compared with each other: L1 speakers of Dutch and Spanish, Dutch learners of Spanish (DLS), and Spanish learners of Dutch (SLD). All participants were raised in a monolingual setting, with either Dutch or Castilian Spanish as their L1. None of them reported any speech language disorders. **Table 1** summarizes further details about the sample that are relevant to the experiment.[1] The L1 speakers of Dutch were students of Tilburg University participating for course credit; none of them spoke Spanish. The L1 speakers of Spanish were students or employees of the Escuela Oficial de Idiomas and the Complutense University of Madrid who participated voluntarily or students at Tilburg University; none of them spoke Dutch. The DLS were Tilburg University students who followed a Spanish course at the university's language center or had learned Spanish before and students from the Spanish Bachelor or Master program at the University of Groningen. The SLD were students or teachers of the Dutch program at the Escuela Oficial de Idiomas, the Complutense University of Madrid, or the University of Groningen. All L2 learners participated on a voluntary basis.[2]

Table 1 *Participant characteristics*

L1	L2	Gender (f/m)	*M* age (*SD*)	Proficiency	*N*
Dutch	–	15/11	23.23 (3.56)	native	26
Spanish	–	13/6	25.32 (7.67)	native	19
Dutch	Spanish	16/5	21.38 (2.20)	less proficient	21
Dutch	Spanish	15/4	26.05 (11.94)	proficient	19
Spanish	Dutch	12/7	22.32 (1.86)	less proficient	19
Spanish	Dutch	17/3	23.60 (6.02)	proficient	20

The L2 groups were further split into two subgroups to allow for additional analyses. In order to do this, an assessment of the participants' L2 proficiency level was made according to the CEFR

(Council of Europe, 2001), which distinguishes between six different proficiency levels (A1, A2, B1, B2, C1, C2) that range from a beginner to a highly proficient speaker. Participants were assigned a proficiency level corresponding to the level of the last course they had successfully completed. Because it was difficult to obtain sufficient participants of each proficiency level to perform meaningful statistical analyses, the participants were split into two groups, both consisting of approximately the same number of participants: the less proficient speaker group (with a language proficiency level of < A1, A1, or A2) and the more proficient speaker group (with a language proficiency level of B1, B2, or C1), see **Table 2**. Although the data set contains a gender bias, with more female (N = 88) than male (N = 36) participants, it is not expected that this will influence the results, as men and women are not known to produce different types of pitch accent distributions, and gender-related F0 effects are controlled for by using equivalent rectangular bandwidth (ERB; see below) as a dependent measure. Lastly, L1 speakers are compared with proficient L2 speakers to investigate bidirectional prosodic transfer. Only the proficient L2 speakers were selected for this analysis to optimally approximate the conditions of Mennen (2004), whose participants were highly advanced L2 speakers.

Table 2 *Distribution of Spanish and Dutch L2 learners by proficiency level*

Spanish learners of Dutch (SLD)		Dutch learners of Spanish (DLS)	
Less proficient	More proficient	Less proficient	More proficient
< A1 = 6	B1 = 9	< A1 = 4	B1 = 6
A1 = 8	B2 = 7	A1 = 9	B2 = 5
A2 = 5	C1 = 4	A2 = 8	C1 = 8

2.4.2 Materials

The study used a picture-naming task designed to elicit intonation patterns in contexts with varying information status. Participants were shown sequences of different objects in various colors and, as the pictures of the objects appeared from left to right, they were asked to name them. All of the objects and colors depicted in both the Dutch and the Spanish experiment had two-syllable names, and only pictures of common and unambiguous objects were used. Furthermore, the pictures mostly elicited voiced sound segments in the target languages in order to facilitate the prosodic analysis at a later stage. The fourth picture of the sequence was always the target image (e.g., a blue donkey) and by varying the preceding objects and their colors, the description of the participant had to coincide with one of four possible types of information status, which were previously used in Swerts et al. (2002) and Krahmer and Swerts (2001):

⇒ Contrastive/Contrastive (CC), in which both the first and the second word of the NP are unmentioned in the preceding descriptions of the pictures in the sequence (e.g., *blue donkey*, preceded by *pink broom*);

⇒ Given/Contrastive (GC), in which the second word of the NP contrasts with the second word of the preceding NP, but the first one is the same (e.g., *blue donkey*, preceded by *blue broom*);

⇒ Contrastive/Given (CG), in which the first word is new in that list, but the second word corresponds (only) with the second word in the description of the preceding picture (e.g., *blue donkey*, preceded by *red donkey*);

⇒ Given/Given (GG), in which both the first and second word are used to describe the preceding picture, but not any other in the sequence (e.g., *blue donkey*, preceded by *blue donkey*).

Note that in Spanish the first word of the NP is the noun, which is followed by the adjective, while in Dutch this order is reversed. This entails that the new/given distributions do not correspond to either a noun or an adjective in both languages. The four types of information status are exemplified in **Table 3** and a sample experimental sequence is illustrated in **Figure 1**.

Table 3 *Examples of the four context conditions (second word is the target)*

Context	Dutch	Spanish
CC	roze bezem, blauwe ezel	guante verde, burro rojo
GC	blauwe bezem, blauwe ezel	burro verde, burro rojo
CG	rode ezel, blauwe ezel	globo rojo, burro rojo
GG	blauwe ezel, blauwe ezel	burro rojo, burro rojo

Note. CC = Contrastive/Contrastive; GC = Given/Contrastive; CG = Contrastive/ Given; GG = Given/Given.

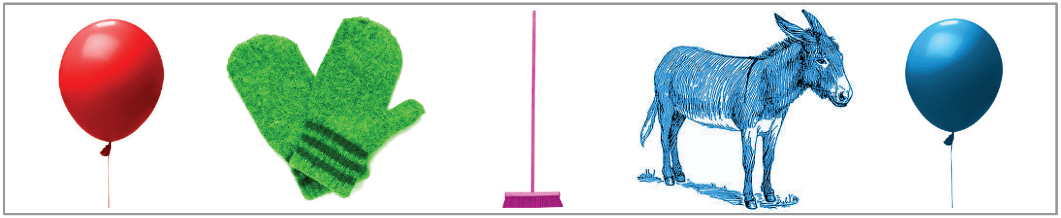


Figure 1. Example of an experimental item, corresponding to a Dutch target noun phrase (fourth object) in a Contrastive/Contrast context.

Four different objects were used in the study; for Dutch: *ezel* ([ezəl], "donkey"), *wanten* ([wɑntən], "mittens"), *ballon* ([bɑlɔn], "balloon"), *bezem* ([bezəm], "broom"); for Spanish: *burro* ([buʀo], "donkey"), *llave* ([ʎaβe], "key"), *guante* ([ɡ̊wante], "glove"), *globo* ([ɡ̊loβo], "balloon"). These objects could have four different colors; for Dutch:[3] *groen* ([xʀun], "green"), *rood* ([ʀot], "red"), *blauw* ([blɑu], "blue"), *roze* ([ʀɔzə], "pink"); for Spanish: *rojo* ([ʀoxo], "red"), *negro* ([neɡ̊ɾo], "black"), *verde* ([beɾde], "green"), *rosa* ([ʀosa], "pink"). This resulted in 16 target items, whose order was randomized for the experiment (4 types of information status × 4 different objects). Every experimental item was alternated with either a filler item in which the participants had to calculate the answer to a simple sum and pronounce it out loud or an item eliciting a sentence with corrective focus of the type *No, it is a* BLUE *donkey*. This made it more difficult for the participants to guess the objective of the study and ensured that they would not perceive the experiment as one long enumeration and would not revert to a flat intonation because of the monotony of the task. It also served as a reset of the discourse status, so that the beginning of each sequence of descriptions could be considered as all new information.

2.4.3 Procedure

The experimental sessions were performed individually and took approximately 15 minutes for the participants who only performed the experiment in one language and 30 minutes for those who also performed the experiment in the L2. The speech recordings, made with the Audacity software (Audacity, Version 2.0.2, 2012) and the internal microphone of an Apple MacBook Pro, occurred in a closed-off (where possible, soundproof) room. Participants were instructed to seat themselves comfortably in front of the laptop, which ensured that they all performed the experiment at a roughly equal distance from the microphone. The task was preceded by a short practice session, in which participants were shown all the objects and colors and asked to describe them, one practice item with corrective focus, and one filler item, so that participants could get used to the type of task and stimuli. In the case of the L2 learners, the practice session provided the researcher the opportunity to ensure that participants knew the words for the object names and colors in the L2, enabling a smooth description of each object. When participants used a non-target word to describe the object or its color, the experimenter corrected them. This type of mistake only occurred during the practice session and therefore did not influence the descriptions of the experimental items. In the rare occasion that participants produced disfluent speech during an experimental item, the researcher would ask them to repeat their description of all the pictures in the list. Participants also filled in a questionnaire concerning their personal information and the languages that they speak to ensure that all participants fulfilled the requirements of each participant group in age, nationality, and L1/L2 and that none of them were restricted in any way when performing the experiment (for instance by colorblindness). Due to time constraints, other extra-linguistic factors, such as the learners' attitudes toward their languages, learning strategies, and motivation, were not taken into account in the current study. The L2 learner groups performed the experiment in one trial that consisted of two conditions, one eliciting speech in the L1 and the other in the L2. To control for possible learning effects, half of the participants first performed the L1 condition and the other half started with the condition in their L2.[4]

2.4.4 Prosodic analysis

Acoustic analyses were performed with Praat (Boersma & Weenink, 2014, version 5.3.68). The target NPs were extracted from the original wave file and segmented into syllables. For Spanish, this was done based on the rules stated in the *Nueva gramática de la lengua española* (Real Academia Española, 2009) and for Dutch on the rules stated in Van der Hulst (1984). Then, the relative difference between the accented and the unaccented words of the NP was calculated. As it has been suggested (Shue et al., 2010) that boundary tones—that is, "the tone— either high or low—associated with the end of an intonational phrase" (Ladd, 2008, p. 80)—may influence the perception of prominence, it is crucial that they be analyzed separately from the pitch accents. Thus, it is not the difference in maximum pitch between the two words of the NP that is measured, but the difference between the maximum pitch values measured within the accented syllable of each word. Consequently, the boundary tone (which in the current data set falls on the last syllable of the NP) is unable to distort the data on pitch accents.

As shown in **Figure 2** on the next page, the maximum pitch value of the vowel in the accent-bearing syllable of both words of the NP was measured. By subtracting the highest F0 value found within the accent-bearing syllable of the first word (A1; "bu" in *burro rojo*, "red donkey") from the highest F0 value found within the accent-bearing syllable of the second word (A2; "ro" in *burro rojo*, "red donkey"), the relative difference between the accented syllables was calculated (i.e., A2 – A1 = pitch accent difference score). This entails that if the second word is pronounced more prominently than the first (e.g., *red* DONKEY), this results in a positive difference score; while an NP in which the first word receives the main prominence (e.g., RED *donkey*) generates a negative difference value. Because one of the aims of the current study is to use an objective acoustic measurement instead of relying on subjective perception measures or coding interpretations, and because previous work (e.g., Kaland, Avesani, Krahmer, Swerts, & Zappoli, 2014) has shown that difference scores accurately represent prominence differences as they are perceived by L1 listeners, only degrees of prominence are investigated. However, future research might investigate the tonal realization of prominence accents, given that there are different ways of realizing a pitch accent (e.g., a rising accent can be realized as L*+H vs. L+H* vs. L+>H*). To get a more complete picture of L1 to L2 prosodic transfer, we also compare the use of boundary tones of both speaker groups. By subtracting the maximum F0 value of the vowel in the accented syllable of the second word of the NP (A2; "ro" in *burro rojo*, "red donkey") from the highest pitch value measured in the last syllable of the NP (%max; "jo" in *burro rojo*, "red donkey"), the relative difference between the accented syllable of the second word of the NP and the boundary tones was calculated (i.e., %max – A2 = boundary tone difference score).

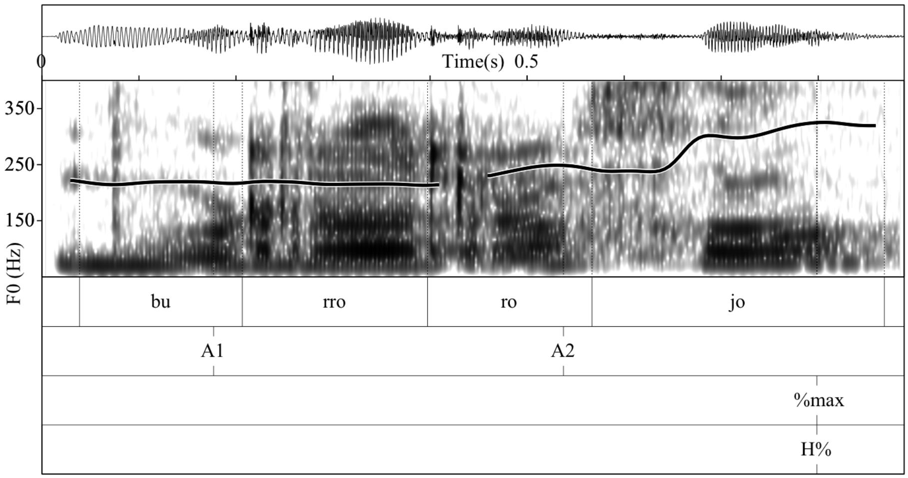


Figure 2. F0 track of a Spanish target noun phrase in a Contrastive/Given context produced by an L1 speaker, including prosodic coding.

To abstract over non-linguistic factors, such as gender or age, measurements were converted to the ERB scale (Glasberg & Moore, 1990), which expresses pitch in a way that is consistent with how speech is perceived by the human ear. A manual correction was performed on items containing an octave jump or other anomalies, such as creaky voice. During the experimental sessions, it became clear that the majority of the participants mistakenly saw the fourth picture of the first experimental sequence as the last one of that list. This was caused by the fact that in the practice session the participants were asked to name the four objects and their colors, which unfortunately led them to think that they would be describing sequences of four objects in the actual experiment, which was not the case. This often resulted in an "end of list intonation," with a falling boundary tone at the end of the description of the fourth picture in the first of our 16 experimental sequences. Because the data points corresponding to this item (124 data points of the 1,984 total data points collected, or 6.25% of the experimental data) could not be directly compared to the other data in which the fourth item correctly was taken as a non-final item and this produced with a rising instead of a falling intonation, the 124 initial data points were excluded from the statistical analysis. All data points corresponding to erroneous descriptions and syllables without an F0 maximum, which were usually due to voiceless or soft speech, were treated as missing values and not taken into account in data analysis (1.18% for the L1 speakers, 1.23% for the L2 data of the L2 speakers, and 2.18% for the L1 data of the L2 speakers), which was performed in IBM SPSS Statistics (IBM Corporation, 2015).

2.5 Results

2.5.1 L1 speech: Setting a baseline (RQ1)

Two repeated-measures analyses of variance (ANOVAs) were performed; one on the pitch accent data and one on the boundary tone data. Both had information status type (CC, GC, CG, GG) as a within-subjects factor, language (L1 Spanish, L1 Dutch) as a between-subjects factor, and the difference scores in ERB as the dependent variable. **Table 4** and **Appendix Figure A1**, available in the appendix at the end of this chapter, summarize the results.

Table 4 *Mean difference score (standard deviation) in equivalent rectangular bandwidth for pitch accents and boundary tones produced by the L1 speakers of Spanish and Dutch, separately by context condition*

Context	Pitch accents		Boundary tones	
	L1 Spanish	L1 Dutch	L1 Spanish	L1 Dutch
CC	0.47 (0.66)	0.11 (0.46)	1.25 (0.55)	0.20 (0.58)
GC	0.46 (0.42)	0.37 (0.55)	1.20 (0.54)	0.23 (0.33)
CG	0.44 (0.53)	–0.40 (0.48)	1.17 (0.73)	0.28 (0.51)
GG	0.46 (0.52)	0.24 (0.46)	1.23 (0.65)	0.32 (0.54)

Note. These data are depicted graphically in the **Appendix** at the end of this chapter. CC = Contrastive/Contrastive; GC = Given/Contrastive; CG = Contrastive/ Given; GG = Given/Given.

With respect to the L1 effect on the production of pitch accents to mark focus, the ANOVA reveals a significant main effect of language, $F(1, 43) = 10.58$, $p < .01$, $\eta_p^2 = .20$, a significant main effect of information status type, $F(3, 129) = 8.99$, $p < .001$, $\eta_p^2 = .17$, as well as a significant interaction between the two, $F(3, 129) = 8.09$, $p < .001$, $\eta_p^2 = .16$. Pairwise comparisons (always Bonferroni-adjusted) explain these effects: While there are no significant differences between the four information status types in Spanish, in Dutch the CG context differs significantly from all other contexts ($p < .001$) and the difference between the CC condition and the GC condition approaches significance ($p = .067$). This means that in Dutch pitch accent distributions reflect focus placement, while in Spanish they do not, as illustrated by the pitch tracks of Dutch and Spanish L1 speakers in the CG and GC conditions in **Figure 3**, on the next page.

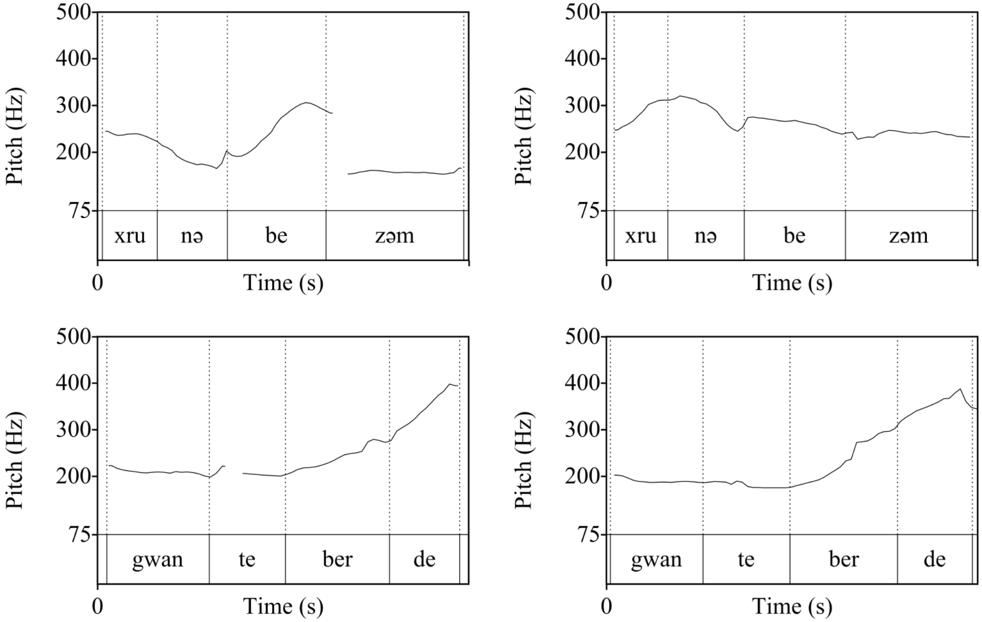


Figure 3. Prototypical F0 tracks of a Dutch target noun phrase (NP; top row) and a Spanish target NP (bottom row) in Given/Contrastive context (left) and in Contrastive/Given context (right) produced by a L1 speaker.

The ANOVA performed on the boundary tone data also reveals a significant main effect of language, $F(1, 43) = 46.02$, $p < .001$, $\eta_p^2 = .52$, but no significant main effect for information status type or a significant interaction effect. Pairwise comparisons show that both speaker groups produced comparable difference scores in all four information status conditions, yet the difference scores by the Spanish L1 speakers are notably higher (on average approximately 1 ERB in all conditions) than those of the Dutch L1 speakers.

2.5.2 L2 speech: Prosodic transfer (RQ2)

This section aims to determine whether prosodic transfer takes place from the L1 to the L2. Therefore, the statistical analysis is limited to the nonnative speech data for the L2 learners. Four repeated-measures ANOVAs were performed (one per L2 group, separately for each prosodic feature) with information status type (CC, GC, CG, GG) as a within-subjects factor, language (L1 Dutch, L1 Spanish, and either L2 Dutch or L2 Spanish) as a between-subjects factor, and the difference scores (either for pitch accents or boundary tones) in ERB as the dependent variable. **Table 5** and **Appendix Figure A2** summarize the results, which are compared to the data produced by the L1 speakers (see **Table 4** and **Appendix Figure A1**). With respect to the SLD, the data on pitch accents show significant main effects for information status type, $F(3, 243) = 13.26$, $p < .001$, $\eta_p^2 = .141$, and language, $F(2, 81) = 7.33$, $p < .01$, $\eta_p^2 = .153$, as well as a significant interaction between language and information status type, $F(6, 243) = 5.08$, $p < .001$, $\eta_p^2 = .111$.

Table 5 *Mean difference score (standard deviation) in equivalent rectangular bandwidth for pitch accents and boundary tones produced in the L2 by the Spanish learners of Dutch (SLD) and the Dutch learners of Spanish (DLS), separately by context condition*

Context	Pitch accents		Boundary tones	
	SLD	DLS	SLD	DLS
CC	0.45 (0.54)	0.31 (0.26)	1.40 (0.77)	0.38 (0.53)
GC	0.61 (0.61)	0.45 (0.28)	1.43 (0.65)	0.43 (0.43)
CG	0.36 (0.69)	–0.09 (0.41)	1.45 (0.69)	0.25 (0.60)
GG	0.64 (0.66)	0.38 (0.31)	1.49 (0.66)	0.39 (0.45)

Note. These data are depicted graphically in the **Appendix** at the end of this chapter. CC = Contrastive/Contrastive; GC = Given/Contrastive; CG = Contrastive/ Given; GG = Given/Given.

Interestingly, pairwise comparisons between language groups reveal that the data produced by the SLD differ significantly from the data produced by the L1 speakers of Dutch ($p < .01$), but not from the data of the L1 speakers of Spanish, suggesting that prosodic transfer is taking place. Pairwise comparisons between information status conditions within the language groups confirm the existence of prosodic transfer: In the Dutch data produced by the SLD, only the CG and GG conditions differ significantly from each other ($p = .016$), while all the other conditions are not significantly different from each other. This suggests that the SLD data are more in line with the Spanish L1 data, in which none of the four information status conditions differ significantly from each other, than with the Dutch L1 data, in which the CG context differs significantly from all other contexts and the difference between the CC and the GC condition, which previously was a trend, has now become a significant effect ($p = .041$), presumably due to the larger sample size.

Analysis of the boundary tones data by the SLD reveal that there is a significant main effect of language, $F(2, 81) = 40.32$, $p < .001$, $\eta_p^2 = .50$, while there is no significant main effect for information status type nor a significant two-way interaction. Pairwise comparisons between language groups indicate a significant difference between the difference scores for boundary tones produced by the SLD and those produced by the L1 speakers of Dutch ($p < .001$), but not between those of the SLD and the L1 speakers of Spanish who do not speak Dutch, suggesting that transfer also occurs in the L2 production of boundary tones. Pairwise comparisons within language groups show that all groups produce comparable difference scores in all information status conditions, but that the difference scores by the SLD lie between those produced by the L1 speakers of Dutch and Spanish. Examination of the mean difference scores produced by the SLD reveals that their values are closer to those of the Spanish L1 speakers than to those of the Dutch L1 speakers.

Next, the analysis[5] of the pitch accent data produced by the DLS reveals a significant main effect of information status type, $F(2.591, 212.484) = 24.70$, $p < .001$, $\eta_p^2 = .23$, and language, $F(2, 82) = 8.41$, $p < .001$, $\eta_p^2 = .17$, with a significant two-way interaction, $F(3.712, 212.484) = 5.11$, $p < .001$, $\eta_p^2 = .11$. Pairwise comparisons between language groups reveal that the pitch accent data produced by the DLS do not differ significantly from the data produced by the L1 speakers of Spanish ($p = .07$), but they do not differ significantly from those produced by the L1 speakers of Dutch either ($p = .06$), implying the existence of transfer. Pairwise comparisons within language groups show that the DLS produce pitch accent patterns that do not correspond to those found in L1 Spanish, but that are similar to those produced by the L1 speakers of Dutch, as the CG condition differs significantly from all other conditions ($p < .001$).

The analysis of the boundary tones produced by the DLS reveals a significant main effect of language, $F(2, 82) = 29.62$, $p < .001$, $\eta_p^2 = .42$, but no significant main effect for information status type or a significant two-way interaction. Pairwise comparisons between language groups indicate a significant difference between the difference scores for boundary tones produced by the DLS and those produced by the L1 speakers of Spanish ($p < .001$), but not between those of the DLS and the L1 speakers of Dutch who do not speak Spanish. Pairwise comparisons within language groups show that all speaker groups produce comparable difference scores in all information status conditions. Furthermore, the difference scores of the DLS fall between those produced by the L1 speakers of Dutch and Spanish, as was the case with the SLD. The mean difference scores produced by the DLS are substantially closer to the values produced by the Dutch L1 speakers than to those of the Spanish L1 speakers, implying the existence of prosodic transfer in the production of boundary tones by the DLS.

2.5.3 L2 speech: Proficiency level (RQ3)

The data were analyzed with two repeated-measures ANOVAs per prosodic feature with information status type (CC, GC, CG, GG) as a within-subjects factor, experiment language as a between-subjects factor (e.g., for Spanish, a distinction is made between Dutch L1 speakers, Spanish L1 speakers, less proficient DLS, and more proficient DLS), and the difference scores in ERB as the dependent variable. **Table 6** and **Appendix Figure A3** summarize the results, which are compared with the L1 data (see **Table 4** and **Appendix Figure A1**).

Table 6 *Mean difference score (standard deviation) in equivalent rectangular bandwidth for pitch accents and boundary tones produced in the L2 by the less and more proficient Dutch learners of Spanish, separately by context condition*

Context	Pitch accents		Boundary tones	
	Less proficient	More proficient	Less proficient	More proficient
CC	0.32 (0.23)	0.29 (0.30)	0.17 (0.30)	0.61 (0.63)
GC	0.43 (0.32)	0.48 (0.24)	0.29 (0.29)	0.58 (0.50)
CG	–0.16 (0.23)	–0.01 (0.54)	0.05 (0.38)	0.48 (0.71)
GG	0.42 (0.31)	0.42 (0.32)	0.24 (0.30)	0.54 (0.53)

Note. These data are depicted graphically in the **Appendix** at the end of this chapter. CC = Contrastive/Contrastive; GC = Given/Contrastive; CG = Contrastive/ Given; GG = Given/Given.

The analysis of the pitch accent data produced by the DLS reveals a significant main effect of information status type, $F(2.592, 209.964) = 29.40$, $p < .001$, $\eta_p^2 = .27$, and language, $F(3, 81) = 5.70$, $p < .01$, $\eta_p^2 = .17$, with a significant two-way interaction, $F(7.776, 209.964) = 3.55$, $p < .001$, $\eta_p^2 = .12$. Pairwise comparisons between language groups reveal no significant difference between the results of the Dutch L1 speakers and those of the DLS, irrespective of their language proficiency, which suggests that transfer occurs in both less proficient and more proficient speech and is therefore not influenced by proficiency level. However, pairwise comparisons within language groups shows that, while the less proficient DLS still produce a difference score in the CG context that is significantly lower than those produced in all other contexts ($p < .01$), the more proficient DLS only produce difference scores that are significantly different when directly comparing the CG and GC contexts and the CG and GG contexts ($p < .01$). This means that the extent to which the learners copy prosodic patterns from their L1 to their L2 diminishes as their proficiency level increases, as illustrated by the pitch tracks of the less and more proficient DLS (**Figure 4** on the next page).

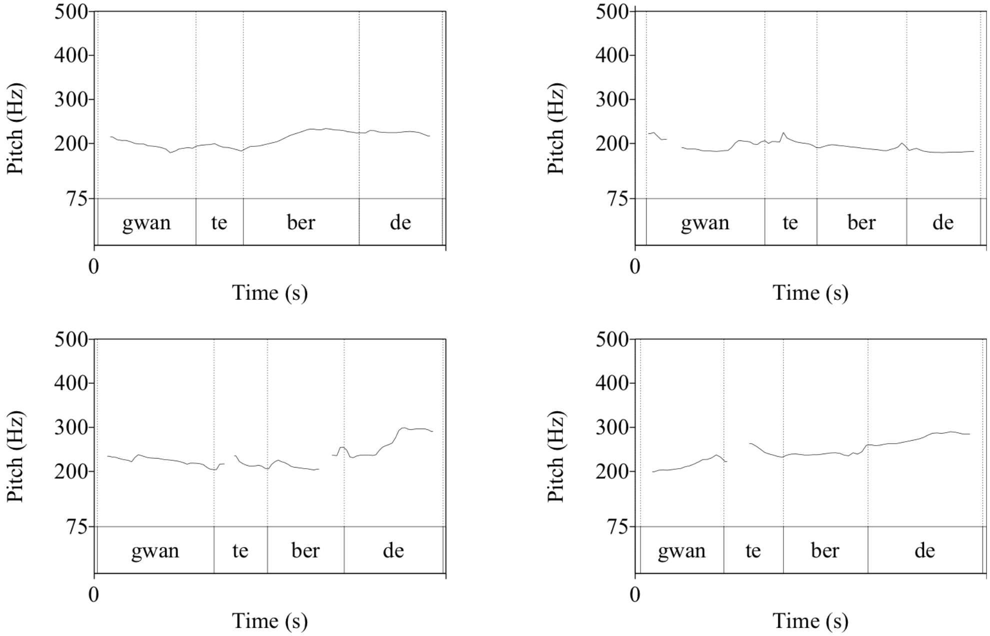


Figure 4. Prototypical F0 tracks of a Spanish target noun phrase in Given/Contrastive context (left) and Contrastive/Given context (right) produced by a less proficient Dutch learner of Spanish (DLS; top row) and a more proficient DLS (bottom row).

The analysis of the boundary tone data produced by the DLS reveals a significant main effect of language, $F(3, 81) = 23.59$, $p < .001$, $\eta_p^2 = .47$, with no significant main effect of information status type or a significant two-way interaction. Pairwise comparisons between language groups show no significant difference between the boundary tones produced by the Dutch L1 speakers and the less or more proficient DLS, while both learner groups do differ significantly from the L1 Spanish group ($p < .001$). Although there was no significant effect involving information status type, inspection of the mean difference scores produced by both learner groups shows that the more proficient DLS approximate the L1 Spanish difference scores substantially more than the less proficient DLS do. Consequently, L2 proficiency might affect the degree of transfer in the production of boundary tones in the L2.

With respect to the pitch accent data produced by the SLD, summarized in **Table 7** and **Appendix Figure A4**, the analysis reveals a significant main effect of information status type, $F(3, 240) = 12.60$, $p < .001$, $\eta_p^2 = .14$, language, $F(3, 80) = 6.83$, $p < .001$, $\eta_p^2 = .20$, and a significant two-way interaction, $F(9, 240) = 3.48$, $p < .001$, $\eta_p^2 = .12$. Pairwise comparisons between language groups reveal no significant difference between the results of the Spanish L1 speakers and those of the SLD, independent of their language proficiency. Yet, the less proficient SLD do differ significantly from the Dutch L1 speakers ($p < .001$), while the more proficient SLD do not, implying that though L2 speech by proficient SLD still shares characteristics with the L1, it also has features that coincide with those of the L2, while this does not hold for the speech by the less proficient SLD. Pairwise comparisons within language groups reveal that the less proficient SLD produce similar difference scores in all contexts, like the L1 speakers of Spanish do. Conversely, the data of the more proficient SLD show that the CG and the GG conditions differ significantly from each other ($p < .05$), suggesting that they are approximating the pitch accent distributions of the L1 speakers of Dutch. This is also demonstrated in the pitch tracks of the less and more proficient SLD (shown in **Figure 5**). Although both language groups start to produce lower difference scores in the CG condition than in all others, the scores of the more proficient SLD are closer to the Dutch L1 scores than the data produced by the less proficient SLD.

Table 7 *Mean difference score (standard deviation) in equivalent rectangular bandwidth for pitch accents and boundary tones produced in the L2 by the less and more proficient Spanish learners of Dutch, separately by context condition*

Context	Pitch accents		Boundary tones	
	Less proficient	More proficient	Less proficient	More proficient
CC	0.62 (0.38)	0.29 (0.62)	1.26 (0.64)	1.53 (0.86)
GC	0.78 (0.38)	0.44 (0.74)	1.32 (0.57)	1.54 (0.71)
CG	0.58 (0.38)	0.16 (0.85)	1.50 (0.72)	1.41 (0.66)
GG	0.76 (0.40)	0.52 (0.83)	1.43 (0.61)	1.54 (0.71)

Note. These data are depicted graphically in the **Appendix** at the end of this chapter. CC = Contrastive/Contrastive; GC = Given/Contrastive; CG = Contrastive/ Given; GG = Given/Given.

The analysis of the boundary tone data by the SLD reveal no significant main effect of information status type or a significant two-way interaction, but does show a significant main effect of language, $F(3, 80) = 26.91$, $p < .001$, $\eta_p^2 = .50$. Pairwise comparisons between language groups show that both the less and more proficient SLD differ significantly from the Dutch L1 speakers, but not from the Spanish L1 speakers. Inspection of the mean boundary tone difference scores produced by the SLD reveals that, instead of producing lower boundary tones to approximate a pattern that is typical for their L2, both proficiency groups produce scores that are even higher than those produced by the Spanish L1 speakers. In summary, the Spanish as well as the Dutch data for both prosodic features suggest that an increase in the proficiency level of the L2 learner yields a more native-like production of intonation in the L2.

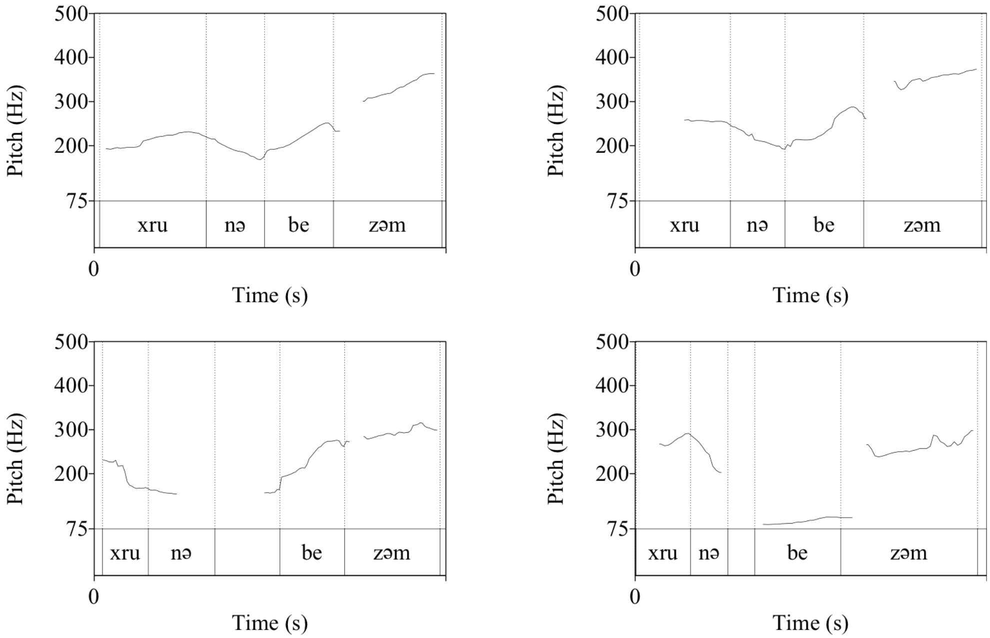


Figure 5. Prototypical F0 tracks of a Dutch target noun phrase in Given/Contrastive context (left) and Contrastive/Given context (right) produced by a less proficient Spanish learner of Dutch (SLD; top row) and a more proficient SLD (bottom row).

2.5.4 L1 speech: Bidirectional transfer (RQ4)

To determine whether the prosodic system of the L2 can also influence the prosodic system of the L1, a comparison is made between the L1 speaker data and the L1 data of the more proficient L2 learners. Four repeated-measures ANOVAs were performed, each with information status type (CC, GC, CG, GG) as a within-subjects factor, language (L1 speech in both languages and L1 speech by the more proficient L2 learners of either Dutch or Spanish) as a between-subjects factor, and the difference scores in ERB as the dependent variable. **Table 8** and **Appendix Figure A5** summarize these results, which are compared to those of the L1 speakers (see **Table 4** and **Appendix Figure A1**). The analysis of the pitch accent data produced by the more proficient SLD in the L1 reveals a significant main effect of information status type, $F(3, 186) = 10.60$, $p < .001$, $\eta_p^2 = .15$, a significant main effect of language, $F(2, 62) = 4.26$, $p < .05$, $\eta_p^2 = .12$, and a significant two-way interaction, $F(6, 186) = 5.06$, $p < .001$, $\eta_p^2 = .14$.

Table 8 *Mean difference score (standard deviation) in equivalent rectangular bandwidth for pitch accents and boundary tones produced in the L1 by the more proficient Spanish learners of Dutch (SLD) and Dutch learners of Spanish (DLS), separately by context condition*

Context	Pitch accents		Boundary tones	
	SLD	DLS	SLD	DLS
CC	0.38 (0.76)	-0.21 (0.33)	1.29 (1.10)	0.13 (0.58)
GC	0.53 (0.70)	0.23 (0.45)	1.13 (0.96)	0.55 (0.66)
CG	0.30 (0.69)	-0.34 (0.44)	1.25 (1.15)	0.31 (0.45)
GG	0.48 (0.76)	-0.02 (0.45)	1.34 (1.06)	0.51 (0.65)

Note. These data are depicted graphically in the **Appendix** at the end of this chapter. CC = Contrastive/Contrastive; GC = Given/Contrastive; CG = Contrastive/ Given; GG = Given/Given.

Pairwise comparisons between language groups show that, even though the SLD do not differ significantly from the Spanish L1 speakers, they also do not differ significantly from the Dutch L1 speakers. Inspection of the mean difference scores produced by the three language groups reveals that, although the SLD produce scores that coincide more with the Spanish L1 speakers in height, they appear to be closer to the pitch accent distributions made by the Dutch L1 speakers: They produce difference scores in the CG condition that are lower than those in all other conditions and difference scores in the GC condition that are higher than those in all other information status conditions. It seems that, where pitch accents are concerned, the SLD speech shows characteristics of both their L1 and their L2. The high boundary tones typical of L1 Spanish were the most difficult feature to adapt to the L2, even for the most proficient SLD; therefore, it is unlikely that those are affected in their L1 speech. The statistical analysis confirms this prediction, as there is no significant main effect of information status type or an interaction effect of information status type; however, there is a significant main effect of language, $F(2, 62) = 16.47$, $p < .001$, η_p^2

= .35. Pairwise comparisons between language groups show that this effect is due to the fact that the difference scores of the SLD in their L1 differ significantly from the difference scores of the Dutch L1 speakers ($p < .001$) but not from those of the Spanish L1 speakers, as is expected.

The analysis of the pitch accent data of the more proficient DLS in the L1 reveals a significant main effect of information status, $F(2.539, 154.878) = 15.83$, $p < .001$, $\eta_p^2 = .21$, and language, $F(2, 61) = 11.95$, $p < .001$, $\eta_p^2 = .28$, and a significant two-way interaction, $F(5.078, 154.878) = 4.73$, $p < .001$, $\eta_p^2 = .13$. Pairwise comparisons between language groups show that the main effect of language is caused by the fact that the DLS data differ significantly from the Spanish L1 data ($p < .001$), but not from the Dutch L1 data, implying that bidirectional transfer does not take place. However, pairwise comparisons within language groups show that the pitch accent distributions of the DLS only partially coincide with the pitch accent distributions of the Dutch L1 speakers: Whereas in L1 Dutch, the CG condition differs significantly from all other conditions ($p < .001$), and the CC condition differs significantly from the GC condition ($p < .05$), the DLS produce difference scores in the CG condition that differ significantly from the GC condition ($p < .01$), but not from the CC and GG conditions. The DLS data also show a significant difference between the CC and the GC condition ($p < .001$).

The analysis of the boundary tone difference scores produced by the DLS shows a significant main effect of information status type, $F(3, 183) = 3.09$, $p < .05$, $\eta_p^2 = .048$, and language, $F(2, 61) = 23.17$, $p < .001$, $\eta_p^2 = .43$, as well as a significant two-way interaction, $F(6, 183) = 2.45$, $p < .05$, $\eta_p^2 = .07$. Pairwise comparisons between language groups reveal that the main effect for language is caused by the fact that the DLS differ significantly from the Spanish L1 speakers ($p < .001$), but not from the Dutch L1 speakers. Pairwise comparisons within language groups reveal that, although there is some variation in the DLS data that does not exist in the L1 data—namely, that the CC condition differs significantly from the GC and GG conditions ($p < .001$)—the pitch accent difference scores produced by the L2 learners do not suggest that either L2 or L1 influence is the cause for this variation.

2.6 Discussion

2.6.1 L1 baseline (RQ1)

It was expected that in Spanish all the information status types would receive comparable and positive difference scores for pitch accents. In Dutch, these difference scores were predicted to depend on information status. Analyses on the L1 data confirmed these predictions. The Spanish L1 speakers produce essentially identical intonation patterns across the different contexts, always placing the main pitch accent on the second word of the NP, as in *guante VERDE* ("green glove"). Conversely, the prominence patterns by the Dutch L1 speakers match the given/contrastive distinctions in the elicited NPs. This is especially clear in the CG context, which receives a negative difference score while the other contexts have markedly higher, positive difference scores, reflecting the intonational difference between, for example, *GROENE bezem* and *groene BEZEM* ("green broom"). Furthermore, the mean difference scores for the GC condition are higher than those of the CC and GG conditions,[6] which reflects the fact that the former conveys narrow focus, while the latter two convey broad focus or complete givenness. In this respect, the data confirm earlier findings (Face, 2002; Hualde, 2005; Krahmer & Swerts, 2001; Nooteboom & Kruyt, 1987; Swerts et al., 2002) and support the predictions formulated for RQ1.

As a result of the comparative nature of the analysis, it was found that the Spanish speakers use substantially higher boundary tones than the Dutch speakers in a comparable context, which has not been reported before. This can partly be explained by our elicitation technique, as the design elicited enumerations of the objects depicted and their colors. It is known that non-final items of a list generally receive a high boundary tone in Spanish (Estebas-Vilaplana & Prieto, 2010). However, the boundary tones produced by some of our speakers reached into the 9.5 ERB, which is rather high. While the use of a high tone to mark non-finality is also used in Dutch (Krahmer & Swerts, 2001), the Dutch L1 speakers produced much lower boundary tones than the Spanish speakers did. Similarly, high boundary tones were found for other Romance languages, such as Italian (Kaland et al., 2014) and Romanian (Swerts, 2007), but it remains an interesting question why our L1 speakers of Spanish exhibit such high boundary tones when L1 speakers of Dutch (and other Germanic languages) use far less extreme ones to convey non-finality.

2.6.2 Prosodic transfer (RQ2)

A comparison of the L1 and L2 data reveals that prosodic transfer occurs both in the use of pitch accents and boundary tones, which is in line with the predictions made based on earlier findings on prosodic transfer in this context (e.g., Gut & Pillai, 2014; Ramírez Verdugo, 2002; Rasier & Hiligsmann, 2009; Swerts & Zerbian, 2010; Zubizarreta & Nava, 2011). As predicted, the L2 data still show characteristics of intonation patterns used in the L1. The results show that prosodic transfer from the L1 to the L2 takes place both for the DLS and the SLD, in the production of both pitch accents and boundary tones. Interestingly, the degree to which transfer takes place appears to be higher in the L2 production of boundary tones than pitch accents. This is not in line with the findings of Swerts and Zerbian, who found that pitch accents were more susceptible to transfer than boundary tones. Furthermore, there are also indications that the L2 learners are acquiring the patterns characteristic of the L2. For instance, the pitch accent difference scores of the SLD show more variation between information status conditions than the L1 Spanish scores; and the scores of the CG condition, although still positive, are much lower than the scores of all other conditions.

2.6.3 Proficiency level (RQ3)

Analyses of the data by the less and more proficient L2 learners reveal that prosodic transfer appears to be influenced by the L2 proficiency level of the speaker. As the L2 proficiency level increases, the speaker produces more native-like intonation patterns. This corroborates earlier claims by Swerts and Zerbian (2010), Zubizarreta and Nava (2011), and Reichle and Birdsong (2014) about the influence of language proficiency on prosodic transfer and by others about linguistic transfer in general. To further demonstrate the modulating effect of proficiency level on the native-like production of intonation, consider the difference scores of the six participants with the lowest L2 proficiency and the six participants with the highest L2 proficiency. Although no statistical significance can be attributed to these results given the small sample size, they clearly show that the proficiency level of the L2 learners influences the use of target-like pitch accent distributions. As shown in **Figure A6**, available in the appendix at the end of this chapter, there is virtually no difference between the data of the least proficient learners and the L1 speakers. The most proficient DLS, however, resemble the Spanish L1 speakers more strongly by producing higher boundary tones and pitch accent difference scores that are closer to those of the L1 Spanish speakers than to those of the least proficient DLS. The most proficient SLD demonstrate the opposite tendency by producing negative difference scores in the CG condition in contrast to all other ones, hereby approaching the contrastive/given distinctions made by Dutch L1 speakers very closely, while the least proficient SLD produce pitch accent difference scores that are typical of their L1. The factor that prevents them from producing native-like intonation in Dutch is their use of very high boundary tones, which even the most proficient SLD continue to produce. This is surprising because the most proficient DLS succeed quite well in approaching the high boundary tones of the L1 speakers of Spanish.

So, where the SLD are able to overcome transfer effects from the L1 to the L2 in the production of pitch accents, but not in the production of boundary tones, the DLS appear to be more successful in the native-like production of boundary tones, and less so in the production of pitch accents. This raises the question whether different factors influence the acquisition process for speakers of different L1s. Are speakers of a plastic language, such as Dutch, less able to adapt to a new way of distributing pitch accents because this is so relevant for communication in the L1? And is it therefore easier for them to inhibit transfer in the production of boundary tones because they are less functional in focus marking in the L1? Conversely, is it easier for speakers of a non-plastic language, such as Spanish, to overcome transfer effects in the distribution of pitch accents, because they do not play a functional role in communication in the L1, but boundary tones do? More research is needed to answer these questions. But the data provide us with an interesting and unexpected paradox, which is worth looking into further. In the interim, prosodic transfer from the L1 to the L2 seems to depend on the proficiency of the L2 learner, both for the SLD and the DLS. As the current data set includes learners from both learning directions, it presents a novel insight in comparison to the study by Swerts and Zerbian (2010), because it shows that the degree of transfer effects and the order in which prosodic features are acquired appears to be influenced by whether the prosodic cue under investigation has communicative value in the L1 or not.

2.6.4 Bidirectional transfer (RQ4)

While no L2–L1 transfer effects were found for boundary tones as predicted, both the DLS and the SLD transfer pitch accent features from their L2 to their L1. Consequently, Mennen's (2004) claims about the existence of prosodic transfer from the L2 to the L1 with respect to peak timing by advanced Dutch learners of Greek are partially corroborated for our learners of Dutch and Spanish. This is surprising because there are several factors that were expected to cause less bidirectional transfer, or none at all. It was assumed that transfer from the L2 to the L1 only occurs when both languages are spoken at a highly advanced level, as Mennen's (2004) participants were highly experienced speakers who used their L2 regularly because they taught it at university level. In that case, bidirectional transfer should not have occurred in our sample, of which the more proficient group included participants from intermediate to near-native level. Further-more, Mennen's participants used their L2 daily in their professional environment, while our participants generally only spoke their L2 in the classroom or with distant relatives. This means that their L1 was usually the dominant language, which would make it difficult for the L2 to influence it. Finally, most of the participants in this study only started learning their L2 during their studies, which were either still ongoing or almost finished. Consequently, the L2 learners' language experience generally did not exceed 6 years, while Mennen's participants all had between 12 and 35 years of L2 experience.

The fact that the current study only partially replicates Mennen's (2004) result may be because her study concerns a phonetic analysis of the intonational form, while our study focused on the functional use of prominence patterns. It can be argued that cross-linguistic influences are realized differently in the concrete realizations of language than in its abstract, phonological dimension. Furthermore, this study investigated context with a high communicative value, while Mennen focused on the timing of a non-final or pre-nuclear rise in declarative intonation, which is a phenomenon that does not necessarily impact the (un)successful communication of a message. It might be that when the use of a certain intonation pattern is relevant to the successful transmission of a message, the necessity of its preservation in the L1 is more evident to L2 learners, as incorrect use may lead to miscommunication or incomprehensibility (Mennen, 2007).[7]

2.7 Conclusion

The current study compared speech data by L1 speakers of Dutch and Spanish to determine the differences between Dutch and Spanish L1 speech in the production of pitch accent distributions to mark focus (RQ1). Consequently, by comparing the L1 data with data elicited from L2 learners of these two languages, we investigated whether prosodic transfer occurs from the L1 to the L2 in the speech by DLS and SLD (RQ2). Follow-up questions regarded whether the proficiency level of the L2 learner influences the degree of prosodic transfer (RQ3) and whether speech produced by proficient L2 learners also showed signs of the L2 influencing the L1 (RQ4). We found that Spanish and Dutch L1 speakers differ in the way they use pitch accent distributions to mark focus in the L1 (RQ1), and this leads to transfer in the speech of both SLD and DLS (RQ2). However, the extent to which transfer takes place is modulated by the proficiency level of the L2 speaker (RQ3) and also appears to be dependent on the prosodic cue under investigation and the communicative value of this prosodic cue in the L1. Unexpectedly, transfer effects from the L2 to the L1 occurred in the production of pitch accents, both in the speech by DLS and SLD (RQ4).

Our research could be extended in a number of ways. First, Spanish is a non-plastic language that employs other means besides the use of pitch accents to mark information status. Although the current experiment elicits semi-spontaneous speech and its design controls for many possible confounds, such as word length and position within the enumeration, it excludes the use of word order changes to indicate information status. In future research, the possibility of a less restrictive design might be investigated, giving speakers more freedom to describe situations as they would in real life (e.g., by means of a map task). Moreover, as the current study shows that transfer, and the degree in which it occurs, is not only dependent on the proficiency level, but possibly also on the L1 of the speaker and the specific prosodic cue under investigation, further research on different combinations of L1s, L2s, and prosodic features with (non-)functional meaning would be highly relevant. Because, in addition to F0, other cues, such as duration and intensity, contribute to the conveyance of prominence (Adams & Munro, 1978), future research might include other prosodic prominence measures.

Furthermore, the present study mainly examines what L2 learners do, yet it lies outside its scope to also investigate *how* they acquire the intonation patterns they exhibit and which methods prove most effective to prevent or reduce prosodic transfer through L2 training. Future research could also examine the exact nature of the acquisition process or pursue the way in which native-like intonation could be taught in an educational setting. While the current study showed that it is difficult for L2 learners to overcome the influence of the L1 on the L2, prosodic features play an important part in successful spoken communication and in passing for a native speaker and therefore might deserve more attention in the classroom as well. Finally, the effect of incorrect production of intonation on intelligibility, comprehensibility, and judgments of nativeness in this particular context could be investigated by means of perception studies.

Notes

[1] As a result of our attempt to balance the language groups for gender and age, there is a larger difference in sample size between the L1 groups than between the L2 groups. However, Levene's tests of equality of error variances examining the difference scores in the four information status conditions reveal that the assumption of equal variances is not violated in the two L1 groups (*F*s < 1.03).

[2] Controlling for knowledge of other Romance languages proved impossible in the Dutch group, as French is obligatory at the highest levels in the Dutch secondary educational system. As our Dutch natives were all university students, we assumed they had completed either the highest or the second highest level of secondary education. Therefore, all the Dutch participants had a (minimal) background in French. Of all the Dutch participants in our study (N = 66), two spoke Italian, but pairwise comparisons showed that their data did not differ significantly from the data of the other participants in their language group. Similarly, all Spanish participants had some knowledge of a West-Germanic language other than Dutch, such as English, as this language is taught in Spanish (high) schools. Of all the Spanish participants in this study (N = 58), 14 participants spoke German, but pairwise comparisons revealed that their data did not

differ significantly from the data of the other participants in their language group.

[3] When the color and the object are combined in one NP, an "e" is added to the adjective. Consequently, all adjectives used in the experiment consisted of two syllables; for example, *groen* became *groe-ne*, *rood* became *ro-de*, and *blauw* became *blau-we*.

[4] Pairwise comparisons showed that the sequence in which the tasks were carried out (L1-L2 vs. L2-L1) had no effect on the results.

[5] A Greenhouse-Geisser correction was always used for the degrees of freedom if the sphericity assumption was violated.

[6] In theory, the GG condition can elicit two utterance types for the third and fourth picture; one where the speaker describes the two identical pictures without additional emphasis and another in which the speaker produces the description of the fourth picture with more emphasis than the third to indicate that they are not simply repeating or correcting their description of the third picture, but that there are two identical pictures and they are describing the second one. While this is a possible confound, it is improbable that it would lead to varying pitch accent distributions. It is unlikely that participants would produce a pitch accent distribution in this context that results in a negative difference score. This view is supported by the fact that our L1 Dutch data show as many positive difference scores in the GG condition as in the CC and GC conditions.

[7] This could also explain why pitch accent distributions appear more difficult to acquire for the DLS than for the SLD. As correct pitch accent distributions are crucial to focus marking in Dutch, they have a highly communicative function, while in Spanish they do not. Therefore, the SLD may have had less difficulty overcoming transfer in their production of pitch accent distributions, as these are less communicatively salient in Spanish.

References

Adams, C., & Munro, R. R. (1978). In search of the acoustic correlates of stress: Fundamental frequency, amplitude, and duration in the connected utterance of some native and non-native speakers of English. *Phonetica*, *35*, 125–156.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, *32*, 233–250.

Audacity (Version 2.0.2). [Computer program]. (2012). Retrieved from http://audacityteam.org.

Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer (Version 5.3.68) [Computer software]. Retrieved from http://www.praat.org

Caspers, J., & Horloza, K. (2012). Intelligibility of non-natively produced Dutch words: Interaction between segmental and suprasegmental errors. *Phonetica*, *69*, 94–107.

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, United Kingdom: Cambridge University Press.

Cutler, A., & Ladd, D. R. (1983). *Prosody—Models and measurements*. Berlin, Germany: Springer.

Cruttenden, A. (1993). The de-accenting and re-accenting of repeated lexical items. In *Proceedings of the ESCA workshop on Prosody*, 16-19.

De Ruiter, L. E. (2015). Information status marking in spontaneous vs. read speech in story-telling tasks: E=vidence from intonation analysis using GToBI. *Journal of Phonetics*, *48*, 29–44.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning 27*(2), 315–330.

Ellis, R. (1994). *The study of second language acquisition*. Oxford, United Kingdom: Oxford University Press.

Estebas-Vilaplana, E., & Prieto, P. (2010). Castilian Spanish intonation. In P. Prieto, & P. Roseano (Eds.), *Transcription of intonation of the Spanish language* (pp. 17–48). Munich, Germany: Lincom Academic.

Face, T. L. (2000). Prosodic manifestations of focus in Spanish. *Southwest Journal of Linguistics*, *19*, 45–62.

Face, T. L. (2002). *Intonational marking of contrastive focus in Madrid Spanish*. Munich, Germany: Lincom Academic.

Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 229–273). Timonium, Maryland, United States: York Press.

Flege, J. E., & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America*, *76*, 708–721.

Flege, J. E., Schirru, C., & MacKay, I. R. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, *40*, 467–491.

Foote, R. (2010). Age of acquisition and proficiency as factors in language production: Agreement in bilinguals. *Bilingualism: Language and Cognition*, *13*, 99–118.

Gabriel, C., & Kireva, E. (2014). Prosodic transfer in learner and contact varieties: Speech rhythm and intonation of Buenos Aires Spanish and L2 Castilian Spanish produced by Italian native speakers. *Studies in Second Language Acquisition*, *36*, 257–281.

Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, *47*, 103–138.

Guion, S. G. (2003). The vowel systems of Quichua–Spanish bilinguals: Age of acquisition effects on the mutual influence of the first and second languages. *Phonetica*, *60*, 98–128.

Gut, U., & Pillai, S. (2014). Prosodic marking of information structure by Malaysian speakers of English. *Studies in Second Language Acquisition*, *36*, 283–302.

Gut, U., Pillai, S., & Mohd Don, Z. (2013). The prosodic marking of intonation status in Malaysian English. *World Englishes*, *32*, 185–197.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, *38*, 201–223.

Henriksen, B., Geeslin, K., & Willis, E. (2010). The development of L2 Spanish intonation during a study abroad immersion program in León, Spain: Global encounters and final boundary movements. *Studies in Hispanic and Lusophone Linguistics*, *3*, 113–162.

Hoot, B. (2012). *Presentational focus in heritage and monolingual Spanish* (Doctoral dissertation). University of Illinois at Chicago, Chicago, United States.

Howell, P., & Kadi-Hanifi, K. (1991). Comparison of prosodic properties between read and spontaneous speech material. *Speech Communication*, *10*, 163–169.

Hualde, J. I. (2005). *The sounds of Spanish*. Cambridge, United Kingdom: Cambridge University Press.

Hualde, J. I. (2009). Unstressed words in Spanish. *Language Sciences*, *31*(2-3), 199-212.

IBM Corporation. (2015). SPSS Statistics (version 22) [Computer software]. Available from http://www.spss.com/software/statistics/

Imai, S., Walley, A. C., & Flege, J. E. (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *Journal of the Acoustical Society of America*, *117*, 896–907.

Jarvis, S. (2000). Semantic and conceptual transfer. *Bilingualism: Language and Cognition*, *3*, 19–21.

Jiang, N. (2004). Semantic transfer and its implications for vocabulary teaching in a second language. *The Modern Language Journal*, *88*, 416–432.

Kaland, C. C. L., Avesani, C., Krahmer, E. J., Swerts, M. G. J., & Zappoli, A. (2014). Adapting to atypical prosody: Contrastive noun phrases in Dutch and Italian. In C.C.L. Kaland, *Prosodic marking of semantic contrasts Do speakers adapt to addressees?* (Doctoral dissertation). Tilburg University, Tilburg, The Netherlands.

Krahmer, E., & Swerts, M. (2001). On the alleged existence of contrastive accents. *Speech Communication*, *34*, 391–405.

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge, United Kingdom: Cambridge University Press.

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*, *32*, 543–563.

Mennen, I. (2007). Phonological and phonetic influences in non-native intonation. In J. Trouvain, & U. Gut (Eds.), *Non-native prosody: Phonetic description and teaching practice* (pp. 53-76). Berlin, Germany: Mouton de Gruyter.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *49*, 285–310.

Nava, E., & Zubizarreta, M. L. (2010). Deconstructing the nuclear stress algorithm: Evidence from second language speech. In N. Erteschik-Shir & L. Rochman (Eds.), *The sound patterns of syntax* (pp. 291–316). Oxford, United Kingdom: Oxford University Press.

Nooteboom, S. G., & Kruyt, J. G. (1987). Accents, focus distribution, and the perceived distribution of given and new information: An experiment. *Journal of the Acoustical Society of America*, *82*, 1512–1524.

Pavlenko, A. (2000). L2 influence on L1 in late bilingualism. *Issues in Applied Linguistics*, *11*, 175–205.

Pavlenko, A., & Jarvis, S. (2002). Bidirectional transfer. *Applied Linguistics*, *23*, 190–214.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation* (Doctoral dissertation). Massachusetts Institute of Technology, Cambridge, Massachusetts, United States.

Ramírez Verdugo, M. D. (2002). Non-native interlanguage intonation systems: A study based on a computerized corpus of Spanish learners of English. *ICAME Journal*, *26*, 115–132.

Rasier, L. (2006). *Prosodie en vreemdetaalverwerving: accentdistributie in het Frans en in het Nederlands als vreemde taal* ('Prosody and second language acquisition: Accent distribution in French and Dutch as a foreign language', in Dutch) (Doctoral dissertation). Université catholique de Louvain, Louvain, Belgium.

Rasier, L., & Hiligsmann, P. (2009). Exploring the L1-L2 relationship in the L2 acquisition of prosody. *Online proceedings of first and second languages: Exploring the relationship in pedagogy-related contexts*, Oxford, United Kingdom. Retrieved from: http://www.education.ox.ca.uk/wordpress/wp-content/uploads/2010/08/RasierHiligsmann.doc.

Real Academia Española. (2009). *Nueva gramática de la lengua española* ('New grammar of the Spanish language', in Spanish). Madrid, Spain: Espasa Calpe.

Reichle, R., & Birdsong, D. (2014). Processing focus structure in L1 and L2 French: L2 proficiency effects on ERPs. *Studies in Second Language Acquisition*, *36*, 535–564.

Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research*, *16*, 135–172.

Shue, Y. L., Shattuck-Hufnagel, S., Iseli, M., Jun, S. A., Veilleux, N., & Alwan, A. (2010). On the acoustic correlates of high and low nuclear pitch accents in American English. *Speech Communication*, *52*, 106–122.

Swerts, M. (2007). Contrast and accent in Dutch and Romanian. *Journal of Phonetics*, *35*, 380–397.

Swerts, M., Krahmer, E., & Avesani, C. (2002). Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics*, *30*, 629–654.

Swerts, M., & Marsi, E. (2012). Prosodic evaluation of accent distributions in spoken news bulletins of Flemish newsreaders. *Journal of the Acoustical Society of America*, *132*, 2616–2624.

Swerts, M., Strangert, E., & Heldner, M. (1996). F0 declination in read-aloud and spontaneous speech. *Proceedings of the Fourth International Conference on Spoken Language Processing*, *3*, 1501–1504.

Swerts, M., & Zerbian, S. (2010). Intonational differences between L1 and L2 English in South Africa. *Phonetica*, *67*, 127–146.

Terken, J. (1984). The distribution of pitch accents in instructions as a function of discourse structure. *Language and Speech*, *27*, 269–289.

Terken, J., & Nooteboom, S. G. (1987). Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes*, *2*, 145–163.

Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, *28*, 1–30.

Ulbrich, C. (2013). German pitches in English: Production and perception of cross-varietal differences in L2. *Bilingualism: Language and Cognition*, *16*, 397–419.

Vallduví, E. (1991). *The informational component*. New York, New York, United States: Garland.

Van der Hulst, H. (1984). *Syllable structure and stress in Dutch*. Dordrecht, The Netherlands: Foris.

Willems, N. (1982). *English intonation from a Dutch point of view*. Dordrecht, The Netherlands: Foris.

Zubizarreta, M. L. (1998). *Prosody, Focus, and Word Order*. Cambridge, Maryland, United States: MIT Press.

Zubizarreta, M. L., & Nava, E. (2011). Encoding discourse-based meaning: Prosody vs. syntax. Implications for second language acquisition. *Lingua*, *121*, 652–669.

Appendix

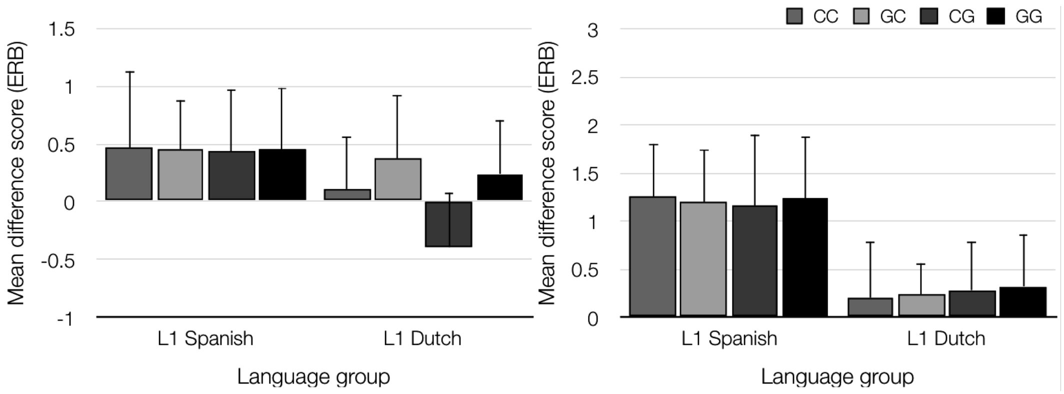


Figure A1. Mean difference scores in ERB for pitch accents (left) and boundary tones (right) produced by L1 speakers of Spanish and Dutch (corresponding to **Table 4**).

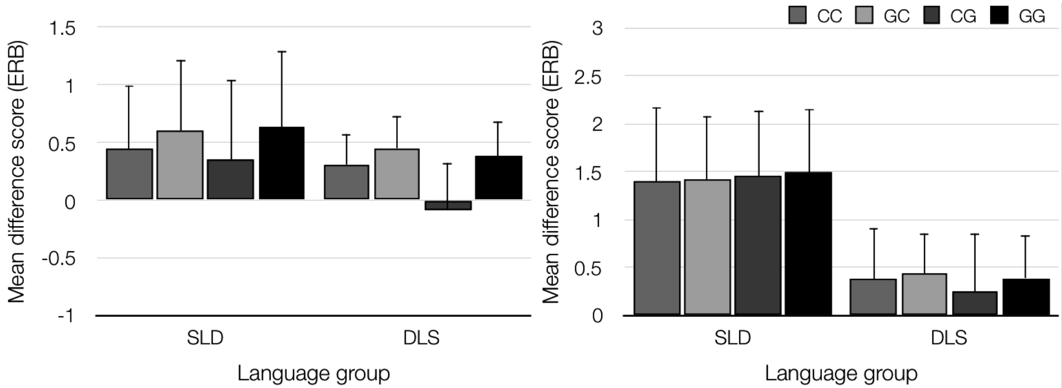


Figure A2. Mean difference scores in ERB for pitch accents (left) and boundary tones (right) produced in the L2 by the SLD and DLS, to be compared with **Figure A1** (corresponding to **Table 5**).

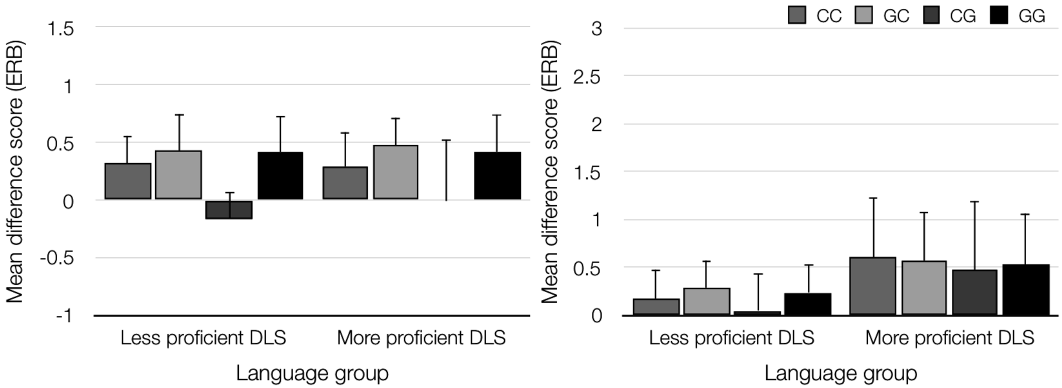


Figure A3. Mean difference scores in ERB for pitch accents (left) and boundary tones (right) produced in the L2 by the less and more proficient DLS, to be compared with **Figure A1** (corresponding to **Table 6**).

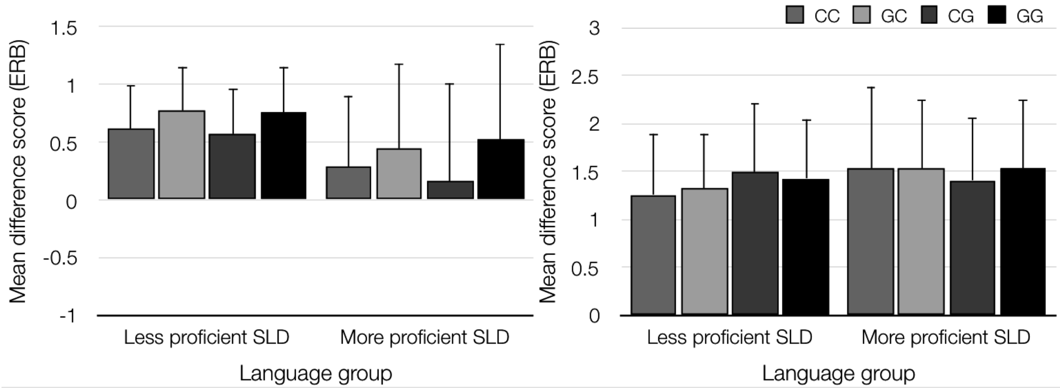


Figure A4. Mean difference scores in ERB for pitch accents (left) and boundary tones (right) produced in the L2 by less and more proficient SLD, to be compared with **Figure A1** (corresponding to **Table 7**).

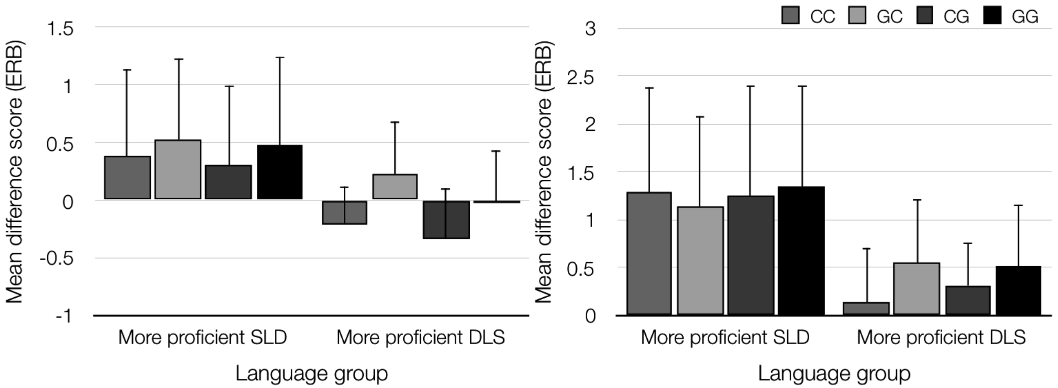


Figure A5. Mean difference scores in ERB for pitch accents (left) and boundary tones (right) produced in the L1 by the more proficient DLS and SLD, to be compared with **Figure 1A** (corresponding to **Table 8**).

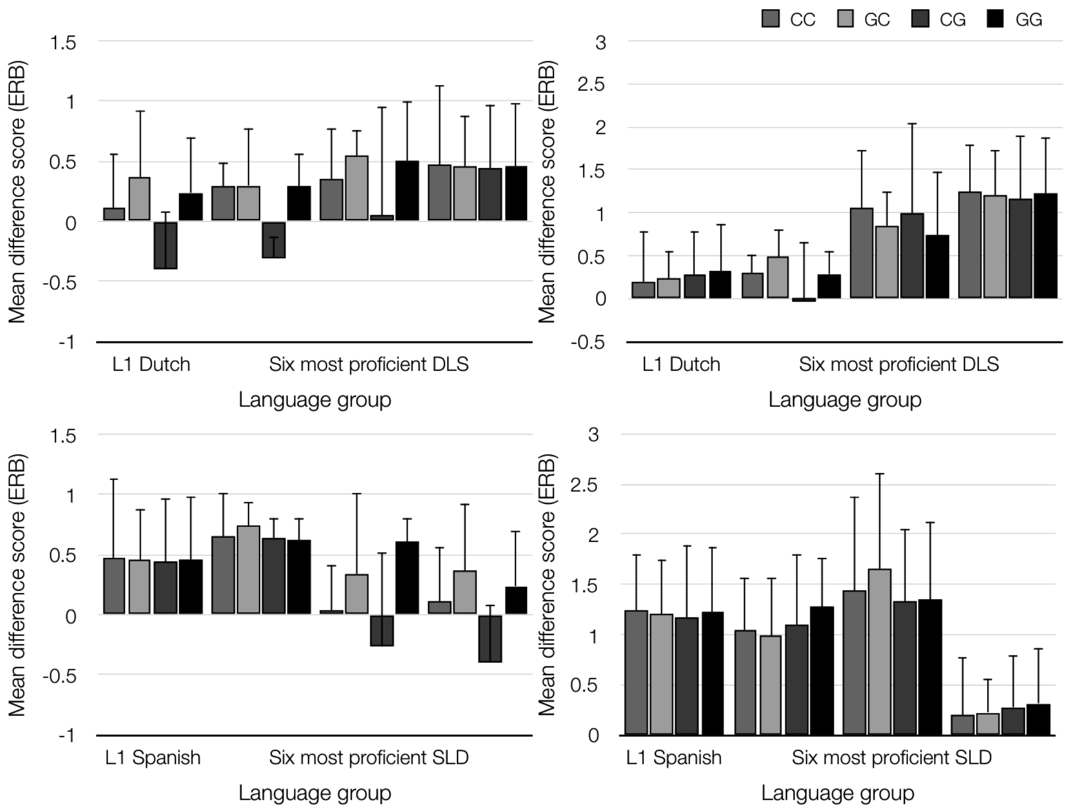


Figure A6. Mean difference scores in ERB for pitch accents (left) and boundary tones (right) produced in the L2 by the six least proficient and the six most proficient DLS (top row), the six least proficient and the six most proficient SLD (bottom row), and L1 speakers of both languages.

3

intonation perception

3

Native Speaker Perceptions of (Non-)Native Prominence Patterns:

Effects of deviance in pitch accent distributions on accentedness, comprehensibility, intelligibility, and nativeness*

Abstract

This research investigates how deviance in focus marking by means of pitch accent distributions by native and non-native speakers affects native speaker perceptions. It shows that non-native speech is rated as less nativelike, more foreign accented and more difficult to understand than native speech, with speakers' proficiency as a modulating factor. Even when controlling for segmental deviance, native listeners could distinguish between two utterances that only differed in whether their focus distribution matched their original elicitation context or not and found the matched utterance more nativelike than the mismatched utterance, based solely on prosodic cues. However, a reaction times experiment reveals that this preference did not influence the processing time of native Dutch utterances with or without prosodic deviance by native listeners. There was a significant difference between reaction times for non-native and native stimuli, but this is most likely due to the slower speech rate of non-native speech.

* **This chapter is based on:** Van Maastricht, L., Krahmer, E., & Swerts, M. (2016). Native speaker perceptions of (non-) native prominence patterns: Effects of deviance in pitch accent distributions on accentedness, comprehensibility, intelligibility, and nativeness. *Speech Communication*, *83*, 21-33.

3.1 Introduction

It is generally assumed that second language (L2) learners have at least two goals: to communicate successfully in the L2 and to approximate native speaker standards as closely as possible. Prior research showed that these goals are often jeopardized by transfer from the native language (L1) to the L2, which can occur in the production of syntactic (e.g., Doğruöz & Backus, 2007; Robertson, 2000; Salamoura & Williams, 2007), semantic (e.g., Jarvis, 2000; Jiang, 2004), segmental (e.g., Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Flege, 1987; Flege & Eefting, 1987; Flege, Schirru, & MacKay, 2003) and suprasegmental features (e.g., Mennen, 2004; Swerts & Zerbian, 2010; Rasier & Hiligsmann, 2009). Transfer effects in the acquisition of segmentals (individual phonemes) of a language and its effect on listeners' perceptions have been studied extensively, more so than transfer effects at the suprasegmental or prosodic level. This makes sense, since the meaning of an utterance is derived mostly from the lexical meaning of the words and the syntactic relations among them, and adequate segmental production is known to be relevant for word recognition. The substitution of one segment for another can change the meaning of a word, for example, the difference between /ˈlik/ ("leek", a vegetable) and /ˈɹik/ ("reek", an unpleasant odour) ensues from the substitution of a segmental that is especially difficult for Japanese learners of English, as their L1 does not have this particular phonemic contrast (Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1997). However, while changing the melodic pattern of a word like "leek" will not change its lexical meaning in intonational (non-tonal) languages such as English, it can change the pragmatic status of the utterance. For instance, changing its boundary tone from a falling one (L%) to a rising one (H%), can transform it from a declarative utterance into a question (Ladd, 1996).

L1 listeners may have judgments about the origins of a L2 speaker based on the rhythm or intonation of their speech (Kolly & Dellwo, 2014). Previous studies have shown that Romance and Germanic languages are distinguishable from one another based solely on prosodic features, by native listeners as young as five days old (Nazzi, Bertoncini, & Mehler, 1998), as well as by adults (Ramus, Dupoux, & Mehler, 2003). Although these studies claim that the rhythmic properties of the language are responsible for correct language discrimination, more recent research suggests that it might actually be the intonational features that allow for correct discrimination (Hagmann & Dellwo, 2014). Either way, these findings show that suprasegmental cues can be equally salient for listeners as segmental ones (cf. Van Els & De Bot, 1987). Unfortunately, prosodic quality is often measured holistically, making it difficult to pinpoint to what extent perceptual differences are related to inadequate functional use of specific prosodic features. While there are studies in which specific prosodic features are investigated (e.g., speech rate, voice quality), it remains unclear if these cues correspond to a specific functional meaning, and if so, what this functional meaning is. For example, while speech rate affects the ease with which L1 listeners processes L2 speech (Kang, 2010; Munro & Derwing, 1998; 2001), the content of the message remains identical, whether it is produced rapidly or slowly. Conversely, the use of pitch accent distributions to mark focus, which is the prosodic cue under investigation in the current study, is expected to change the functional meaning of an utterance, as deviance in its production changes the communicative value of the message and affects how the utterance fits in a larger discourse context. For instance, if speaker A would ask speaker B *Would you like chocolate or strawberry ice cream?,* an answer like *I would like chocolate* ICE CREAM (small capitals indicating emphasis) would be confusing to speaker A, because the emphasis on *ice cream* suggests a choice between different foods (e.g., chocolate ice cream or chocolate cake), instead of an option between two flavours, as was A's intention. However, if speaker B's L1 does not use pitch accents to mark focus to the same degree as many Germanic languages do (Rasier & Hiligsmann, 2009; Swerts, 2007; Swerts, Krahmer, & Avesani, 2002; Chapter 2 of this dissertation), they might not be able to produce a pitch accent on a prefinal element and deaccent the final element, as is required to make the appropriate contrast: *I would like* CHOCOLATE *ice cream* (e.g., Wennerstrom, 1994, 1998; Pickering, 2001, 2004). Therefore, incorrect placement of pitch accents can obstruct communication between speakers.

Thus, the innovative character of the current study partly lies in its focus on a specific function of prosody, that is, the use of pitch accent distributions to mark focus. Earlier work on the effect of pitch accent placement on word recognition and focus interpretation in the L1 shows that this particular prosodic cue is relevant for the speed with which L1 listeners process speech (e.g., Terken & Nooteboom, 1987), but the question remains how important this is when suprasegmental deviance is combined with deviance in other cues in the L2, such as segmental, syntactic and semantic features. Intuitively, one might assume that the prosodic cues continue to affect processing speed, but it could also be the case that when other error types are abundant, suprasegmental errors no longer play an important role. In the present study, we perform a series of studies in order to investigate the effect of deviance in the production of pitch accent distributions to mark focus on L1 listeners' perceptions concerning L1 and L2 speech, with a focus on Spanish and Dutch.

3.2 Theoretical background

3.2.1 L2 perception: Accentedness, comprehensibility, intelligibility, and nativeness

Prior work on the perception of L2 speech by L1 listeners is often based on different measures and contexts and reveals diverging results. However, there are four concepts that frequently reappear: ACCENTEDNESS,[1] COMPREHENSIBILITY, INTELLIGIBILITY, and NATIVENESS. For sake of clarity, a short definition of each of these terms is provided. For the first three we adopt the definitions of Derwing and Munro (2005, p. 385): ACCENTEDNESS refers to "how much an L2 accent differs from the variety of speech commonly spoken in the community", COMPREHENSIBILITY concerns "the listener's perception of the degree of difficulty encountered when trying to understand an utterance", and INTELLIGIBILITY "the extent to which the speaker's intended utterance is actually understood by a listener". These last two terms are especially ambiguous, as they are used interchangeably in previous work (e.g., Anderson-Hsieh, Johnson, & Koehler, 1992; Anderson-Hsieh & Koehler, 1988; Edmunds, 2009; Hahn, 2004). NATIVENESS is understood to be "the degree to which a speaker sounds like a native speaker of a particular language" (Edmunds, 2009, p. 2), but is sometimes also referred to as naturalness (e.g., Maassen & Povel, 1984) or taken to be the opposite of accentedness. In other words, accented speech corresponds to L2 speech and L1 speech is generally considered to be unaccented (Edmunds, 2009; Schairer, 1992).

There are several studies reporting effects of suprasegmental cues on one or more of these four constructs for different language pairs (e.g., Trofimovich & Baker, 2006; Boula de Mareüil & Vieru-Dimulescu, 2006; Edmunds, 2009; Hahn, 2004; Grover, Jamieson, & Dobrovolsky, 1987). Generally, these studies conclude that deviance in the production of prosodic features influences listeners' perceptions of accentedness and comprehensibility, and to a lesser extent nativeness. Hahn (2004) shows that deviance in phrasal stress placement in English also affects intelligibility by measuring students' processing and recall capacities of three versions of a Korean lecture. Others compare the influence of suprasegmental versus segmental features on these constructs (e.g., Van Heuven, Kruyt, & De Vries, 1981; Ulbrich, 2013; Anderson-Hshieh et al., 1992; Magen, 1998; Field, 2005). Generally, these studies report that suprasegmental deviance is more relevant for accentedness, comprehensibility and intelligibility than segmental deviance. Studies on the influence of suprasegmental cues on intelligibility performed outside the field of L2 acquisition corroborate the findings reported for L2 learners: Cutler (1976) showed that new words that receive a pitch accent are recognized faster than new words that are unaccented in English. Terken and Nooteboom (1987) investigated whether the opposite holds (in Dutch) for words that are given in the discourse, indeed showing that these are recognized faster when not pitch accented, while new words are recognized faster when they are. This suggests that phrasal pitch accents used to mark focus influence intelligibility measured though processing time (cf. also Cutler, Dahan, & Van Donselaar, 1997; Van Donselaar, Koster, & Cutler, 2005; Akker & Cutler, 2003).

Caspers and Horloza (2012) report different findings: in their study on the influence of segmental and lexical stress errors on intelligibility, L1 Dutch performed transcription and reaction times (RTs) tasks with stimuli produced by L1 speakers of Dutch, or French and Chinese L2 speakers, and rated the stimuli on accentedness. The experiments reveal that transcription accuracy and RTs are most influenced by a combination of segmental and suprasegmental errors, and the contribution of either of them to intelligibility is roughly equal. This contradicts earlier findings that prosodic cues are more relevant to intelligibility than segmental cues, at least for lexical stress. Similarly, Ulbrich (2013) concludes that both segmental and suprasegmental factors influence perceived accentedness, without one clearly contributing more than the other.

There are a few exceptions to those studies that are limited to one particular construct. Trofimovich and Isaacs (2012) show that accentedness ratings for French learners of English are more related to rhythmic, segmental, and syllable structure accuracy, while comprehensibility ratings are more influenced by grammatical accuracy and lexical richness. However, the only prosodic cue that has a functional meaning in their study, word stress placement, appears to affect both accentedness and comprehensibility ratings equally, with Pearson correlation coefficients of −.78 and −.76, respectively. In Munro and Derwing (1995, 1999) and Derwing and Munro (1997), accentedness, comprehensibility, and intelligibility, but not nativeness, are used as factors in a correlation analysis, showing that although L1 listeners are relatively quick to mark speech as accented, their comprehensibility judgments are less extreme, and the results of intelligibility tests demonstrate that even L2 speech that is rated as accented, is still intelligible. Additionally, they report that prosodic errors contributed more to accentedness and comprehensibility than segmental errors did, while for intelligibility the reverse was found to be true. This suggests that different cues (segmental vs. prosodic) matter for different constructs (accentedness, comprehensibility or intelligibility). Accentedness and comprehensibility were measured on a nine-point scale and intelligibility through transcriptions (Derwing & Munro, 1997; Munro & Derwing 1995) and RTs (Munro & Derwing, 1999). Additionally, intonation was rated on a holistic 9-point scale (1 = native-like, 9 = not at all native-like) by naïve listeners. As intonation can serve a whole range of functions, focussing on one specific feature might enable us to examine its effect on communication between L1 and L2 speakers.

Additionally, most previous studies used speech samples elicited by means of an oral reading task or explicitly instructed speakers to produce a certain pitch accent distribution (e.g., Anderson-Hsieh et al., 1992; Anderson-Hsieh & Koehler, 1988; Caspers & Horloza, 2012; Cutler, 1976; Derwing, Munro, & Wiebe, 1998; Field, 2005; Grover et al., 1987; Hahn, 2004; Magen, 1998; Munro, 1995; Munro & Derwing, 1999; Swerts & Vroomen, 2015; Terken & Nooteboom, 1987). This may be problematic because spontaneous and read-aloud speech have significantly different prosodic characteristics (e.g., Blaauw, 1994; Dellwo, Leemann, & Kolly, 2015; De Ruiter, 2015; Howell & Kadi-Hanifi, 1991; Laan, 1997; Swerts, Strangert, & Heldner, 1996). Since L2 learners supposedly intend to use their L2 in communication, the use of (semi-)spontaneous speech samples that approximate the prosodic features used in real-life communication is important in studies focussing on the effect of prosodic deviance on L1 listeners' perceptions. This is especially relevant in the context of focus marking by means of pitch accents, as De Ruiter (2015) reports that there are significant differences between spontaneous and read speech in this respect. In spontaneous speech, L1 speakers of German always tend to accent new referents (i.e., they use rises in the melodic pattern to make words sound more prominent), but they do not always deaccent[2] given referents. Conversely, read speech is characterized by deaccentuation of given elements and the use of pitch accents for new elements. Additionally, different types of pitch accents are used in the different speaking modes. De Ruiter concludes that reading intonation is markedly different from spontaneous intonation, and that this has important consequences for the methods used in language acquisition research.

In summary, various studies investigated the influence of suprasegmental deviance on accentedness, comprehensibility, nativeness, and/or intelligibility, with mixed results. Most of them only investigate one or two of these constructs at a time, and focus on different prosodic features (e.g., intonation patterns, lexical stress placement, phrasal stress placement) that have different functions (continuity marking, word recognition, focus marking). Therefore, comparing them remains a difficult task, especially with the aim of drawing reliable conclusions concerning the effect of prosodic deviance on real-life communication. In the next section, we will elaborate on the use of pitch accent distributions to mark focus in Dutch and Spanish in order to outline the prosodic feature under investigation in the current research.

3.2.2 Pitch accent distributions to mark focus in Dutch and Spanish

Prosodic marking of focus is an aspect of speech that is difficult to acquire for L2 learners, especially when their L1 is a Romance language and their L2 a West-Germanic one, or vice versa (Rasier, 2006;

Chapter 2 of this dissertation). This is because the languages are typologically different in their use of prosodic properties to highlight important information in a discourse (Cruttenden, 1993; Vallduví, 1991): West-Germanic languages are traditionally taken to be plastic languages in which intonation is used to mark the information status of certain elements by means of pitch accents. In Dutch, this tends to entail that both new and contrastive information are accented, while given information is usually deaccented (Krahmer & Swerts, 2001; Nooteboom & Kruyt, 1987; Terken, 1984). Conversely, Spanish, a Romance language, is considered to be a non-plastic language in which the information status of an element is commonly reflected by the position of the element in the sentence and the fact that the nuclear pitch accent is placed at that, usually fixed, position (Zubizarreta, 1998). In Spanish, broad and narrow focus utterances generally have the same pitch pattern (Face, 2002): the nuclear accent tends to be placed on the last content word of the intonational phrase (Hualde, 2005, p. 257). To place narrow focus on a pre-final element, the standard word order of the utterance can be modified (cf. Hoot, 2012) to ensure that the nuclear accent falls on the focused element. Example (1), a Dutch adaption from Face (2000: 46), illustrates this: In (1a) the direct object *boek* receives a pitch accent and is therefore most prominent. In order for the direct object to receive the nuclear accent in Spanish, the canonical word order in which the direct object appears directly after the verb, is adapted to one which it is placed at the right periphery of the utterance (1b).

(1a) *Juan gaf het* BOEK *aan María.*
John gave the BOOK to Mary.
'John gave the BOOK to Mary'.

(1b) *Juan le dio a María el* LIBRO.
John to+her gave to Mary the BOOK.
'John gave the BOOK to Mary'.

Resulting from this difference, a comparison of speech produced by L1 Dutch and L1 Spanish in a context with a fixed word order (i.e., a noun phrase or NP), but with varying information status types, showed diverging intonation patterns for the two groups, see Chapter 2 of this dissertation. In the NPs by Spanish L1 speakers the pitch accent was always placed on the last word of the NP (e.g., *guante* ROJO, "red glove"). In L1 Dutch, however, prominence depended on which element was new or contrasting in the context: for example, RODE *wanten* ("red mittens") if the adjective was new or contrastive, and *rode* WANTEN if the noun was.

These differences between L1 Dutch and Spanish result in prosodic transfer in NPs produced by L2 learners of both languages. Chapter 2 of this dissertation compared L1 speakers with proficient and less proficient learners of both languages and report that transfer occurs from the L1 to the L2 in both learning directions but is modulated by the proficiency level of the speaker. These findings corroborate prior studies on prosodic transfer in focus marking (e.g., Gut & Pillai, 2014; Gut, Pillai, & Mohd Don, 2013; Rasier & Hiligsmann, 2009; Ramírez Verdugo, 2002; Swerts & Zerbian, 2010). And although previous work has addressed the effect of phrasal prominence on oral proficiency (e.g., Kang, Rubin, & Pickering, 2010; Wennerstrom, 1998) and perceived accentedness (Kang, 2010), it remains unclear whether the use of prominence (by means of pitch accent placement) is functionally relevant in the context of focus marking. Additionally, the question remains whether these specific characteristics of L2 speech that deviate from L1 norms are relevant for the intelligibility of L2 speech, and more specifically whether possible perceptual effects are also relevant for communication between L1 and L2 speakers.

3.3 The current study

The research question addressed in this study is:

RQ1 How does deviance in pitch accent distributions that are used to mark focus influence the constructs of accentedness, comprehensibility, nativeness and intelligibility as perceived by native listeners?

Additionally, we ask for each construct whether the proficiency level of the speaker modulates these possible effects. Study 1 is a rating task, performed to confirm the robustness of the effects found in prior work. This also functions as a baseline for the other two experiments, as it will show whether the spontaneous speech samples used in our experiments, produced by the different language groups described in Chapter 2 of this dissertation, are sufficiently diverse to be suitable for rating and whether the differences between these three groups are discernible for L1 listeners. Study 2 is a preference task performed to assess whether the findings of the rating task are due to segmental or prosodic factors: By controlling for segmental variation and only manipulating the pitch accent distributions in L1 and L2 speech, this task is set up to determine whether L1 listeners prefer utterances in which the pitch accent distributions match the focus structure to utterances in which the pitch accent distributions do not. It will also show whether the utterances produced spontaneously by L1 and L2 speakers are diverse enough for L1 listeners to perceive differences between the three language groups (i.e., L1 speakers, more proficient L2 speakers, and less proficient L2 speakers). Study 3 is a reaction time experiment performed using the same stimuli as the preference task, to assess the effects of deviance in pitch accent distributions on intelligibility. Although intelligibility is often operationalized in previous research by means of a transcription task (e.g., Derwing & Munro, 1997; Munro & Derwing, 1999), the prosodic features under investigation here require a different approach: a correct transcription of a speech fragment demonstrates that a listener has heard all its individual words correctly, yet it does not show whether he has grasped the pragmatic meaning and understands how the utterance fits in the larger discourse context. Previous studies investigating the effect of deviance in pitch accent distributions on intelligibility in L1 and L2 speech (e.g., Cutler, 1976; Swerts & Vroomen; 2015) have shown that measuring the time a listener needs to process a sentence is a reliable way to measure intelligibility in this context. Additionally, given that there is little lexical variation between the stimuli, this makes them unsuitable for a transcription paradigm. We predict the following:

Study 1: Rating task

⇒ Based on previous work concerning L1 perception of the accentedness, comprehensibility, and nativeness of L2 speech (e.g., Derwing & Munro, 1997; Edmunds, 2009; Munro, 1995; Munro & Derwing, 1995; Munro & Derwing, 1999; Van Els & De Bot, 1987), we predict that Dutch L1 listeners are able to distinguish between speech produced by L1 Dutch and Spanish learners of Dutch concerning accentedness, comprehensibility, and nativeness. It is expected that the effects found for non-spontaneous speech stimuli can be reproduced for semi-spontaneous speech fragments.

⇒ Although very few studies incorporate proficiency level as a between-subjects factor in their design when investigating one of these three concepts (see Trofimovich & Baker, 2006, for a notable exception), it is expected that L1 speech is rated as less accented than proficient L2 speech, which in turn is expected to be rated as less accented than less proficient L2 speech. For the comprehensibility and nativeness ratings the opposite tendency is expected. Previous work by Edmunds (2009) on nativeness suggests that suprasegmental cues affect nativeness ratings by L1 listeners. As some have seen nativeness and accentedness as two extremes (e.g., Schairer, 1992), it is predicted that L1 listeners can differentiate between L1 and L2 speech in this respect. To our knowledge, no studies exist that report on the effect of proficiency level on nativeness ratings, but it is predicted that accentedness and nativeness ratings will follow similar patterns.

Study 2: Preference task

⇒ Prior work suggests that L1 listeners are capable of distinguishing prosodically correct speech from prosodically incorrect speech, both in the L1 (Van Heuven, Kruyt & De Vries, 1981) and L2 (Edmunds, 2009). However, less is known about the differentiation within L2 proficiency levels. We predict that participants will be able to distinguish between utterances in which the pitch accent distributions match the focal context of the original elicitation condition and those which do not, wherever the difference between both utterance types is perceivable, that is, in speech produced by L1 speakers of Dutch and proficient Spanish learners of Dutch.

Study 3: Reaction time task

⇒ As we have seen, earlier work on intelligibility paints a mixed picture concerning the processing effects of prosodic deviance. As all studies suggest that there is at least some effect of prosodic deviance on intelligibility, we predict that L1 listeners are faster to process speech by L1 speakers than speech by L2 speakers, irrespective of proficiency level. Furthermore, we predict that RTs for proficient L2 speakers will be faster than those of less proficient L2 speakers, and that L1 listeners will process stimuli produced by L1 speech with a correct pitch accent distribution faster than L1 stimuli with an incorrect pitch accent distribution.

Concluding, combining accentedness, comprehensibility, intelligibility, and nativeness as dependent variables in one study, as well as investigating the relationship between them, has to our knowledge, not been done before. It enables us to draw meaningful conclusions concerning the overall effects of prosodic deviance in L2 speech on L1 perception. The innovative character of the current study further lies in that it investigates how deviance in pitch accents to mark focus affects these constructs by means of a design that uses semi-spontaneous speech samples, a natural way of manipulating prosodic accuracy, focuses on one specific suprasegmental cue, used in a specific functional context, and enables comparison between subjective ratings and objective processing measures. In this way, we hope to demonstrate the relevance of pitch accent distributions to mark focus on different aspects of L1-L2 speaker communication.

3.4 Study 1: Rating task

3.4.1 Method

Listeners

41 Dutch monolinguals participated in the experiment (25 women, age M = 21.20, SD = 1.92, and 16 men, age M = 22.38, SD = 2.06), all students of Communication and Information Sciences at Tilburg University. They participated for partial course credit and none spoke Spanish as a L2.

Materials

The experiment was presented in the online questionnaire format LimeSurvey (2014, version 1.92+). All utterances used in the study were of the type '*het is* + determiner + adjective + noun' ("it is" + determiner + adjective + noun). They consisted of semi-spontaneous speech, which was previously elicited and recorded during a production study (Chapter 2 of this dissertation) in which L1 speakers of Dutch and Spanish learners of Dutch were asked to read aloud the question that appeared on the screen and answer this question basing their answer on the picture also depicted on the screen. **Figure 1** shows the individual pictures used in the elicitation task including the corresponding question (From left to right, top to bottom: "Is this a blue donkey?", "Are these red mittens?", "Is this a pink balloon?", "Is this a green broom?"). The sentence-final NP in the answer of the participants thus referred to differently coloured objects, and all objects and colours had two-syllable names, as shown in (2) - (5), which are the answers to the questions that accompanied the pictures. In all utterances, the NP received broad focus.

(2) *Ja, het is een blauwe ezel*
("Yes, it is a blue donkey")

(3) *Ja, het is een groene bezem*
("Yes, it is a green broom")

(4) *Ja, het is een roze ballon*
("Yes, it is a pink balloon")

(5) *Ja, het zijn rode wanten*
("Yes, they are red mittens")

Figure 1. Stimuli used in Chapter 2 of this dissertation to elicit semi-spontaneous utterances of the type "Yes, this is a blue donkey".

Speakers

The utterances were produced by three types of speakers: four L1 speakers of Dutch (L1), four proficient Spanish learners of Dutch (L2+) and four less proficient Spanish learners of Dutch (L2-). An assessment of the speakers' L2 proficiency level was made according to the Common European Framework of Reference for Languages (Council of Europe, 2001), which distinguishes between 6 different proficiency levels (A1, A2, B1, B2, C1, C2). Speakers were assigned a proficiency level corresponding to the level of the last course they had successfully completed. For the current study a distinction was made between less proficient speakers (proficiency level ≤ A2) and proficient speakers (proficiency level ≥ B2). **Table 1** contains an overview of the details about the speaker sample that are relevant to the experiment.

Table 1 *Characteristics of the speakers used in the rating task*

L1	L2	Gender	Age	Proficiency	*N*
Dutch	-	Male	19 and 23 years	Native	2
Dutch	-	Female	22 and 23 years	Native	2
Spanish	Dutch	Male	45 and 69 years	Proficient	2
Spanish	Dutch	Female	31 and 48 years	Proficient	2
Spanish	Dutch	Male	33 and 34 years	Less proficient	2
Spanish	Dutch	Female	22 and 24 years	Less Proficient	2

Procedure

While the experiment was performed on an online platform, the experimental sessions always took place in a sound proof cubicle to ensure that participants would focus on their task. Sessions were performed individually and took approximately 15 minutes. Participants were instructed to listen to an utterance after which they were asked to indicate their opinion on a nine-point semantic differential scale. The following instructions and corresponding scales for accentedness (6), comprehensibility (7), and nativeness (8) were presented to the participants. Each scale was measured in a separate block of questions, that is, participants first listened to the stimuli to assess their accentedness, then they listened again to rate their comprehensibility, and finally, they evaluated them based on their nativeness. These are the English translations of the original Dutch sentences.

(6) Indicate to which extent the speaker you heard has a foreign accent
No foreign accent – Very strong foreign accent

(7) Indicate to which degree the speaker you heard is easy/difficult to understand
Incomprehensible – Very easy to understand

(8) Indicate to which degree you think that this speaker sounds like a native speaker of Dutch
Does not sound like a native speaker at all – Sounds like a native speaker

The task was preceded by a basic questionnaire to ensure that all participants fulfilled the requirements of each participant group in age, nationality, L1/L2, and that none of them were restricted in any way when performing the experiment (e.g., by visual, auditory or cognitive impairments). As a result of these requirements, three of the, originally 44, participants were excluded: two because they spoke Spanish as a L2, and one participant was excluded due to self-reported hearing difficulties. Consequently, analyses were performed on the data of the 41 remaining participants, who provided ratings for all experimental items. As each of the speakers produced 12 stimuli (3 constructs × 4 different objects), the experiment contained 144 stimuli in total. The experimental items were presented in a random order, both within each block of questions and across participants.

3.4.2 Results

Before performing statistical tests, the accentedness ratings were transformed to reflect the same direction of effect as expected for comprehensibility and nativeness by means of the formula [NewRating = −1 × Rating + 10]. A Generalized Linear Mixed Model analysis was then performed in IBM SPSS Statistics 22 (IBM, 2015) with Speaker Group (3 levels: L1, L2+, and L2-) and Construct (3 levels: accentedness, comprehensibility, and nativeness) as fixed factors. Item, Speaker, and Listener were set as random factors, and Rating was the response variable. Following Barr, Levy, Scheepers, & Tily (2013), models were fitted with a maximal random effects structure, including random intercepts, and random slopes for both fixed effects and their interaction. **Figure 2** summarizes the results. The analysis revealed a significant main effect of Speaker Group, $F(2, 5895) = 28.17$, $p < .001$) and Construct, $F(2, 5895) = 1067.08$, $p < .001$), as well as a significant interaction, $F(4, 5895) = 391.57$, $p < .001$). Pairwise comparisons using the Bonferroni method were performed to further investigate the effects of Speaker Group within each construct.

ACCENTEDNESS. Pairwise comparisons between speaker groups within this construct reveal that there is a significant difference between the three levels (L1 vs. L2+: $p < .001$, L2+ vs. L2-: $p < .001$, L1 vs. L2-: $p < .001$). For the transformed accentedness ratings the effect is positive, in the sense that as the proficiency level of the speaker increases, the corresponding accentedness ratings increase as well (L1: $M = 8.8$, $SD = .8$, L2+: $M = 5.1$, $SD = 2.6$, L2-: $M = 2.4$, $SD = 1.5$).

NATIVENESS. Pairwise comparisons between speaker groups within this construct reveal that there is a significant difference between the three levels (L1 vs. L2+: $p < .001$, L2+ vs. L2-: $p < .001$, L1 vs. L2-: $p < .001$). The effect for nativeness is also positive: as the proficiency level of the speaker increases, the nativeness ratings assigned to the corresponding utterances likewise become higher (L1: $M = 8.8$, $SD = .7$, L2+: $M = 4.9$, $SD = 2.5$, L2-: $M = 2.2$, $SD = 1.5$).

COMPREHENSIBILITY. For comprehensibility, a positive effect is seen as well: as the proficiency level of the speaker increases, the comprehensibility ratings also increase (L1: $M = 8.5$, $SD = 1.1$, L2+: $M = 7.4$, $SD = 2.1$, L2-: $M = 5.9$, $SD = 2.7$). However, while the pairwise comparisons between the speaker groups within this construct reveal a significant difference between the L1 and the L2- groups (p = .001), and a marginally significant difference between the L2+ and L2- groups (p = .058), there is no significant difference between the L1 and L2+ groups (p = .130).

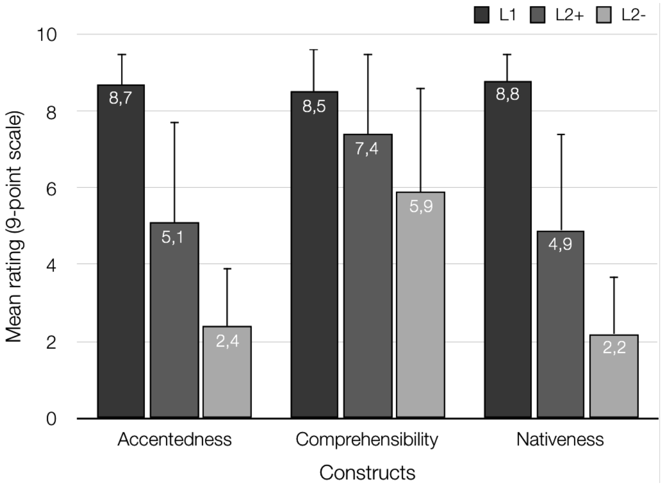


Figure 2. Mean ratings (1-9 scale) and standard deviations by L1 speakers of Dutch for accentedness, comprehensibility and nativeness for stimuli produced by L1 speakers of Dutch, proficient L2 speakers of Dutch and less proficient L2 speakers of Dutch.

A Spearman's Rho correlation analysis was performed to examine the relation between the different construct ratings. It reveals a significant positive correlation between accentedness ratings and comprehensibility ratings, r = .70, p (one-tailed) < .001, and nativeness ratings, r = .92, p (one-tailed) < .001, respectively, in such a way that as accentedness ratings increase, the comprehensibility and nativeness ratings increase as well. Additionally, the analysis revealed a significant positive correlation between the latter two ratings, r = −.76, p (one-tailed) < .001.

3.4.3 Discussion

The results show that L1 listeners of Dutch have clear intuitions concerning the accentedness, comprehensibility and nativeness of L1 Dutch and L2 Dutch produced by L1 speakers of Spanish. They clearly perceive L2 speech as more accented, more difficult to understand and less nativelike than L1 speech. The fact that the accentedness and nativeness constructs yield very similar results, and there is a strong correlation between the two constructs, is not surprising as it has been noted in the Introduction that these have been considered extremities of the same spectrum in previous work (Schairer, 1992). Furthermore, the distinction between the two proficiency levels of the L2 speakers is reflected by the ratings assigned by L1 listeners for these two constructs. It this sense, the intuitions of L1 listeners are very much in line with what is expected and what has been reported in previous work (e.g., Derwing & Munro, 1997; Edmunds, 2009; Munro, 1995; Munro & Derwing, 1995; Munro & Derwing, 1999; Van Els & De Bot, 1987). Furthermore, it shows that the results obtained in previous studies using read aloud or non-spontaneous speech can be replicated with speech stimuli whose prosodic features are closer to speech used in real-life communication between L1 and L2 speakers, confirming both of the predictions made in the Introduction.

Interestingly, the difference between the L2+ group and the other speaker groups is not significant for comprehensibility. This implies that while speech by L1 speakers is significantly easier to comprehend than speech produced by less proficient L2 speakers, the difference in ease of comprehension between the L1 group and the L2+ group, and the two L2 groups respectively, is not as substantial. This is further reflected in the relative difference (Cohen's *d*) between the language groups: regarding the accentedness and the nativeness ratings these are substantially bigger than the relative difference between the language groups for the comprehensibility ratings (Accentedness: L1 vs. L2+ $d = 1.9$, L2+ vs. L2- $d = 1.3$, Comprehensibility: L1 vs. L2+ $d = .7$, L2+ vs. L2- $d = .6$, Nativeness: L1 vs. L2+ $d = 2.1$, L2+ vs. L2- $d = 1.3$). This suggests that L1 listeners are more extreme in their judgments of a foreign accent and perception of non-nativelike speech, than they are concerning the difficulty with which this foreign accented and non-native sounding speech is understood. Despite the fact that L1 listeners clearly hear the difference between the three speaker groups, and intuitively know that certain speech samples sound more foreign, they do not indicate that this influences the difficulty with which they understand the speech substantially. These findings are in line with those of Munro and Derwing (1995), who also report that although they found a correlation between the strength of a foreign accent and comprehensibility, a strong foreign accent does not obligatorily reduce the intelligibility of non-native speech.

Concluding, it can be stated that the robustness of the effects found in earlier work using non-spontaneous speech stimuli can be confirmed for semi-spontaneous speech stimuli as well. Additionally, it is clear that L1 listeners have unambiguous and adequate intuitions concerning the degree of foreign accent in L1 and L2 speech and make corresponding judgments concerning the nativeness of the speaker. Although these judgments are reflected in their comprehensibility ratings, all speaker groups have a mean comprehensibility rating of a 5.9 or higher on a 9-point scale, and can therefore be considered to be relatively easy to understand.

3.5 Study 2: Preference task

3.5.1 Method

Listeners

45 Dutch monolinguals participated in the experiment (32 women, age $M = 24.14$, $SD = 7.72$, and 13 men, age $M = 33.33$, $SD = 16.28$), all students of Communication and Information Sciences at Tilburg University receiving partial course credit. None spoke a Romance language as L2, with the exception of French,

which is taught at Dutch high schools. None participated in any of the other tasks reported in this paper.

Materials

The experiment, performed within the online questionnaire format Qualtrics (2014), consisted of a forced-choice task. All of the utterances used as stimuli in the study contained objects and colours with two-syllable names. Like the ones used in the rating task, the items for this study were elicited in a previous task reported in Chapter 2 of this dissertation. In this part of that particular research, L1 and L2 speakers of Dutch were asked to name objects and their colour as they appeared on a screen from left to right (see Figure 5 in section 3.6.1 for an example stimulus), resulting in an enumeration of the type "pink balloon, green broom, red mittens, blue donkey, pink broom". In these sequences of pictures, the third and fourth object and their colour were varied to manipulate the focus structure of the fourth picture of the sequence. For the current task, the descriptions produced by L1 and L2 speakers of Dutch corresponding to the third and fourth picture were extracted from the enumeration. These utterances were used as stimuli in the present task and could be one of two types: 1) the noun of the second NP could be in narrow focus (e.g., *blauwe bezem, blauwe [ezel]*$_F$, "blue broom, blue donkey"), or 2) the adjective of the second NP could be in narrow focus (e.g., *rode ezel, [blauwe]*$_F$ *ezel,* "red donkey, blue donkey"). **Table 2** shows the stimuli presented to the participants during the preference task in both focus conditions.

Table 2 *Stimuli presented to the participants during the preference task*

Noun is focused		Adjective is focused	
Matching utterance	Mismatching utterance	Matching utterance	Mismatching utterance
Blauwe bezem, blauwe [ezel]$_F$ ("Blue broom, blue donkey")	*Blauwe bezem, [blauwe]*$_F$ *ezel* ("Blue broom, blue donkey")	*Rode ezel, [blauwe]*$_F$ *ezel* ("Red donkey, blue donkey")	*Rode ezel, blauwe [ezel]*$_F$ ("Red donkey, blue donkey")
Groene wanten, groene [bezem]$_F$ ("Green mittens, green broom")	*Groene wanten, [groene]*$_F$ *bezem* ("Green mittens, green broom")	*Rode bezem, [groene]*$_F$ *bezem* ("Red broom, green broom")	*Rode bezem, groene [bezem]*$_F$ ("Red broom, green broom")
Rode ballon, rode [wanten]$_F$ ("Red balloon, red mittens")	*Rode ballon, [rode]*$_F$ *wanten* ("Red balloon, red mittens")	*Groene wanten, [rode]*$_F$ *wanten* ("Green mittens, red mittens")	*Groene wanten, rode [wanten]*$_F$ ("Green mittens, red mittens")

Two types of identical utterances were presented to the participants: for one utterance, the original recording was used in which naturally the focus distribution matched the original context in which it was elicited (the matching condition, e.g., "red donkey, [blue]$_F$ donkey"). For the other utterance, this was not the case, that is, the part of the utterance that concerned the second NP was elicited in a context where focus was on the noun but is presented in a context where focus was on the adjective, and vice versa (the mismatching condition, e.g., "red donkey, blue [donkey]$_F$"). In other words, in these utterances the focus distribution did not match the original elicitation condition, which has consequences for the pitch accent distributions used by the different speaker groups, as shown in the next subsection. Crucially, the prosodic features of the semi- spontaneous fragments that were elicited for Chapter 2 of this dissertation were never manipulated in any way. All stimuli were created by combining a NP in which the focus distribution matches the original elicitation condition with the same NP in which the focus distribution does not match the original elicitation condition.

Speakers

The utterances used in this task were produced by three types of speakers: L1 speakers of Dutch (L1), proficient Spanish speakers of Dutch (L2+), and less proficient Spanish speakers of Dutch (L2-, see the previous Speakers section for more details on the speaker groups). One would predict that the difference between the matching and the mismatching condition is clearly perceivable in utterances produced by L1 speakers of Dutch, who naturally accent the focused word of the NP and deaccent the given word, but less so in utterances by proficient L2 learners, who succeed in placing the main pitch accent on the focused word, but do not deaccent the given word of the NP, and almost indistinguishable in utterances produced by less proficient learners of Dutch, who use almost identical prosodic patterns in both conditions, in which they place the main pitch accent on the last word of the NP, irrespective of its focal status, as is the norm in their L1 (see **Figure 3**). For more information on the prosodic characteristics used by prototypical speakers of the three language groups and corresponding pitch tracks, see Chapter 2 of this dissertation. In this design, segmental quality is controlled for, because a comparison is made between two utterances that are identical in content, produced by the same speaker. By manipulating deviance in pitch accents in this way, its effect on perceived naturalness by L1 listeners could be observed. The design resulted in 72 different items (2 contexts × 3 objects × 12 speakers). The order of the two utterances within the experimental item, as well as the order in which the items were presented to the participants, was randomized.

Procedure

Although the experiment was performed on an online platform, the experimental sessions always occurred in a sound proof cubicle to make sure participants were not distracted during the task. Sessions were performed individually and took approximately 20 minutes. Participants were instructed to listen to two utterances, followed by the question "Which of the two utterances sounds most natural to you?" The task was preceded by a short block of questions to verify that all participants met the requirements of each participant group in age, nationality, L1/L2, and that none of them were restricted by hearing impairments. The data of 4 participants were excluded from analysis, because they spoke either Italian or Portuguese as a L2. Consequently, analyses were performed on the data of the 41 remaining participants.

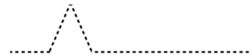
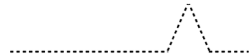
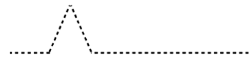
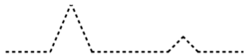
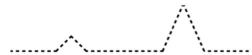
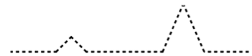
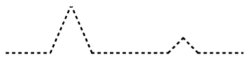
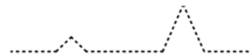
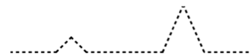
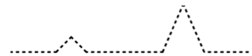
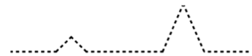
Speaker group	Focussed element	Match between focus distribution and elicitation context	Pitch distribution within NP
			blauwe bezem
L1 Dutch	Adjective	Match	
	Adjective	Mismatch	
	Noun	Match	
	Noun	Mismatch	
Proficient Spanish learners of Dutch	Adjective	Match	
	Adjective	Mismatch	
	Noun	Match	
	Noun	Mismatch	
Less proficient Spanish learners of Dutch	Adjective	Match	
	Adjective	Mismatch	
	Noun	Match	
	Noun	Mismatch	

Figure 3. Stylized pitch contours of the target NP *blauwe bezem* ("blue broom") produced by native speakers of Dutch and more and less proficient Spanish learners of Dutch in all possible conditions.

3.5.2 Results

Following Heck, Thomas and Tabata (2013), we used a Generalized Linear Mixed Model to perform a binary logistic regression analysis in SPSS Statistics 22 (IBM, 2015) with Speaker Group (3 levels: L1, L2+, and L2-) as a fixed factor and Preference (binary: either a '1' for the matching utterance or a '0' for the mismatching utterance) as the response variable. Listener, Speaker, and Item were once again set as random factors. Following Barr et al. (2013), models were fitted with a maximal random effects structure, including random intercepts, and random slopes for the fixed effect. **Figure 4** summarizes the results.

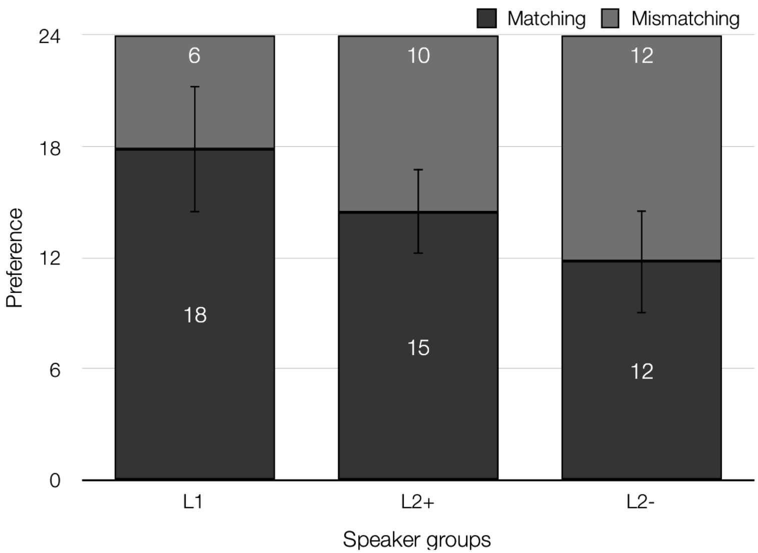


Figure 4. Mean number of preferences (max. 24) and standard deviations by L1 Dutch for utterances with matching or mismatching pitch accent distributions, for stimuli produced by L1 speakers of Dutch, proficient L2 speakers of Dutch, and less proficient L2 speakers of Dutch.

The analysis revealed a significant main effect of Speaker Group, $F(2, 2949) = 33.89$, $p < .001$. Pairwise comparisons between speaker groups reveal that all levels differ significantly from one another ($p < .001$, for all). Additionally, inspection of the mean preference scores (mean number of times the listener preferred the matching utterance to the mismatching utterance) per speaker group reveal that while the utterance in which the focus placement matches the pitch accent distributions is chosen substantially more often when the items are produced by L1 (M = 17.9, SD = 3.4) and L2+ (M = 14.9, SD = 2.3) speakers, L1 listeners have no preference for either of the two utterance types when presented with speech by L2- learners (M = 11.8, SD = 2.8). Note that the maximum number of preferred utterances is 24 per speaker group (4 speakers per proficiency level × 3 items × 2 focus contexts). An almost perfect balance between a preference for utterances in which focus placement and pitch accent distributions match on one hand, and utterances in which this is not the case on the other hand, demonstrates that listeners choose an utterance at random for L2- learners.

3.5.3 Discussion

The results of the preference experiment can be interpreted on several levels. First, the fact that L1 listeners prefer the utterance in which focal status and pitch accent distributions are matching over its mismatching equivalent for the items that are produced by L1 speakers or proficient L2 speakers shows that L1 listeners are able to distinguish between the two utterances based solely on their prosodic characteristics, whenever these characteristics are produced correctly. From this it follows that the difference between prosodically accurate and prosodically inaccurate speech is intelligible even in semi-spontaneous speech. Thus, it does not appear to be true that the difference between prosodically accurate and prosodically inaccurate speech can only be perceived in manipulated or non-spontaneous speech. We therefore assume that these items are suitable for our RT task to examine the effect of deviance in pitch accent distributions on intelligibility as well.

Second, the results show that the preference of L1 listeners for the utterance in which the pitch accent distribution matches the original elicitation context becomes substantially bigger, whenever the difference between the two utterances is audibly distinguishable, as predicted in the Introduction. This suggests that listeners find speech in which focus marking by means of pitch accent distributions is done correctly significantly more natural than speech in which this is not the case (less proficient L2 speech), or less so (proficient L2 speech). This leads to the assumption that L1 listeners perceive deviance in pitch accent distributions both in L1 and L2 speech and feel this affects the naturalness (nativeness) of speech. Although the results presented until now may seem intuitive, the question that remains to be answered is whether this particular phenomenon, that is, deviance in pitch accent distributions to mark focus, ultimately affects the processing of L1 and L2 speech by L1 listeners, which we will investigate in a RT experiment.

3.6 Study 3: Reaction time task

3.6.1 Method

Listeners

The same participants that performed the rating task (but not the preference task), continued to perform the RT task.

Materials

The materials used in the RT experiment are the picture sequences used in Chapter 2 of this dissertation to elicit object names and colours (see Section 3.5.1). The objects and colours depicted all had two-syllable names and only included common and unambiguous objects. Four different objects are used in the task: *ezel* ("donkey"), *wanten* ("mittens"), *bezem* ("broom"), and *ballon* ("balloon") as a practice item. These objects could have four different colours: *groen* ("green"), *rood* ("red"), *blauw* ("blue"), and *roze* ("pink"), see **Figure 5** for an example of a list of objects. These objects and colours were used to form sequences of five pictures that were shown to the participants in combination with utterances naming the object and its colour. By varying the third picture, the information status of the fourth picture was manipulated: half of the stimuli corresponded to a description of the fourth picture that consisted of a NP with focus on the noun (*blauwe bezem, blauwe [ezel]*$_F$, "blue broom, blue donkey"), and the other half of the stimuli corresponded to a description of the fourth picture that consisted of a NP in which the adjective was focussed (*rode ezel, [blauwe]*$_F$ *ezel*, "red donkey, blue donkey").

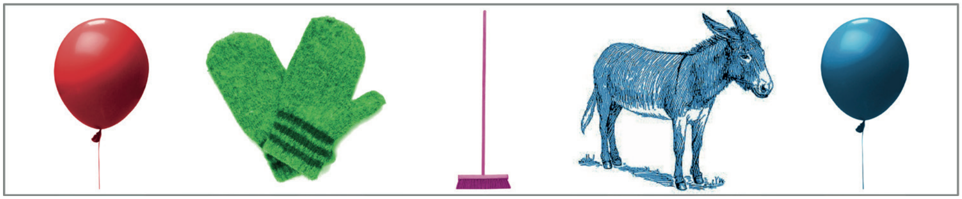


Figure 5. Example of a sequence of experimental objects. Further explanation is found in the text.

Speakers

The manipulation in this task consisted of the fact that different types of speakers gave a description of the objects, which led to different pitch accent distributions in the descriptive utterances. Utterances produced by the same three speaker types as those in the preference task were used for the RT task: utterances by L1 speakers of Dutch (L1+), utterances by proficient Spanish L2 learners of Dutch (L2+), and utterances by less proficient Spanish L2 learners of Dutch (L2-), see Section 3.5.1 for an explanation of the pitch accent distributions they typically produce. An additional speaker group was added to investigate whether deviance in pitch accent distributions to mark focus also affects intelligibility in L1 speech: utterances produced by this L1- group naturally were not foreign accented but had been manipulated so that they contained a contextually inappropriate pitch accent distribution. These incorrect pitch accent distributions were obtained in the same way as those for the preference task, that is, by combining a particular focus context (either the first word of the NP is focused, or the second) with an utterance in which the pitch accent distribution corresponds to the opposite focus context. For instance, a sequence in which the third picture depicts a red donkey and the fourth (target) picture a blue donkey normally elicits a description of the type "red donkey, BLUE donkey", but in the L1- condition these pictures were combined with an utterance of the type "red donkey, blue DONKEY". As Spanish and Dutch differ in the way they mark narrow or contrastive focus but use similar pitch accent distributions to mark broad focus (see Chapter 2 of this dissertation), only the description of the target object was manipulated, while the descriptions of all other objects in the series were as they were produced in the original broad focus elicitation context.

To ensure that participants could not predict whether the target object would match the description they heard or not, half of the pictures corresponded with the descriptions, while the other half did not, which would balance their affirmative and negative responses (see **Table 3**). This design resulted in a total of 192 different items: 2 true/false contexts (experimental vs. filler condition) × 2 focus conditions (noun vs. adjective) × 3 objects × 16 speakers.

Table 3 *Stimuli presented to the participants during the preference task*

Condition	Noun is focused	Adjective is focused
Picture and utterance correspond (experimental condition)	"*Blauwe bezem, blauwe [ezel]*$_F$" (Blue broom, blue donkey)	"*Rode ezel, [blauwe]*$_F$ *ezel*" (Red donkey, blue donkey)
Picture and utterance do not correspond (filler condition)	"*Groene wanten, groene [bezem]*$_F$" (Green mittens, green broom)	"*Rode bezem, [groene]*$_F$ *bezem*" (Red broom, green broom)

Procedure

During the RT task, participants saw the previously mentioned sequences of objects in different colours appear on a computer screen and heard the corresponding descriptions as the objects appeared. For each fourth object that appeared participants were asked to indicate whether the description they heard through their headphones corresponded to the picture they saw on the screen by pressing either a green (yes) or a red (no) button. They were asked to do this as quickly and as accurately as possible, knowing that they could press a button even before the end of the presented utterance. Pilot studies showed that it was difficult for participants to remain focused and alert during the 45 minutes needed to perform all the RT items. Therefore, the experiment was divided into two parts. Participants performed the entire experiment in two sessions of roughly 20 minutes on consecutive days. To avoid any additional variation, the different tasks were always performed in the same order: during the first session, participants performed the previously reported rating task and the first part of the RT task, and during the second session they performed the second half of the RT task. To keep distractions to a minimum, the auditory descriptions were presented to the participants through headphones, and experimental sessions were performed individually in a sound proof cubicle. The RT task, which was made and presented to the participants using E-prime (version 2.10.353, 2014), was preceded by a short practice session, to familiarize participants with the type of task and stimuli. To minimize learning effects, the order of the experimental items was randomized within each session of the RT task.

3.6.2 Results

In order to investigate the influence of deviance in pitch accent distributions on the RTs of L1 listeners of Dutch, a Generalized Linear Mixed Model analysis was performed in SPSS Statistics 22 (IBM, 2015) with Speaker Group (4 levels: L1+, L1-, L2+, and L2-), and

Focus Condition (2 levels: focus on the noun, and focus on the adjective) as fixed factors, and the RTs in milliseconds as the response variable. As in previous analyses, Item, Speaker and Listener were set as random variables. Following Barr et al. (2013), models were fitted with a maximal random effects structure, including random intercepts, and random slopes for both fixed effects and their interaction. As the conditions in which utterances did not correspond to the picture shown on screen were merely used as fillers, they were excluded from the statistical analysis. Earlier experiments that measured intelligibility in a similar way (e.g., Swerts & Vroomen, 2015), measured RTs from the offset of the utterance played after the stimulus picture was shown, which enables listeners to explicitly not focus on the prosodic cues they receive and only concentrate on the content of the utterance, which would minimize the effect of prosodic cues on RTs. To avoid this, in the current study RTs were measured from the onset of the utterance, which coincided with the appearance of the picture, until the participant pushed a button. This entails that participants were able to react before the stimulus utterance was completed and could therefore use prosodic cues to predict the focal status of the elements of the NP. For example, if they heard *blauwe bezem, GROENE* ... ("blue broom, GREEN ...") it would be safe to assume in Dutch that the next object might also be a broom. RTs that deviated more than two standard deviations from the overall mean were excluded from the analysis, as were RTs corresponding to incorrect responses. This resulted in the total omission of 5.36% of the entire dataset. **Figure 6** summarizes the results.

The analysis revealed a significant main effect of Speaker Group, $F(3, 3758) = 5.28$, $p = .001$, and of Focus Condition, $F(1, 3758) = 9.87$, $p = .002$, and a marginally significant interaction, $F(3, 3758) = 2.43$, $p = .063$. The main effect for Focus condition is caused by the fact that utterances in which the adjective was focused ($M = 858$, $SD = 201$) were processed significantly faster than utterances in which the noun was focused ($M = 907$, $SD = 226$). This is to be expected, as the adjective is always the first word of the Dutch NPs and therefore faster to corroborate for listeners. Pairwise comparisons between language groups reveal that the L1+ group differs significantly from both L2 groups (L1+ vs. L2+: $p = .014$, L1+ vs. L2-: $p = .001$), as does the L1- group (L1- vs. L2+: $p = .014$, L1- vs. L2-: $p = .001$). Interestingly, the RTs for utterances produced by the two L1 speaker groups do not differ significantly from one another neither when collapsing over both focus conditions ($p = .899$), nor when analysing them separately ($p = .782$ when the noun is focused, $p = 1.000$ when the adjective is focused). The same can be said for both L2 groups, which do not differ significantly from each other when Focus Condition is not taken into account ($p = .844$), nor when the adjective or noun is focused ($p = 1.000$, and $p = .522$, respectively).

To control for the lower speech rate of L2 speakers compared to L1 speakers, an additional Generalized Linear Mixed Model analysis was performed. In this analysis, the same fixed and random factors were used as in the initial analysis, but for the response factor the RTs were measured as a proportion of the length of the stimulus (SL) using the formula $[(100 \times RT) / SL]$. The analysis reveals a significant main effect of Focus Condition, $F(1,3757) = 4.37$, $p < .05$, but no significant effect of Language Group, $F(3,3757) = .11$, $p = .957$, nor a significant two-way interaction, $F(3,3757) = .44$, $p = .726$. Inspection of the mean normalized RTs per

Language Group show that none of the groups differ significantly from one another when speech rate is controlled for (L1+: M = 86.6, SD = 20.0; L1-: M = 86.9, SD = 19.2; L2+: M = 86.3, SD = 16.4; L2-: M = 84.5, SD = 15.2). The main effect of Focus Condition is caused by the fact that participants responded significantly faster to stimulus items in which the focus was on the adjective (M = 85.1, SD = 17.6) than to stimulus items in which the focus was on the noun (M = 87.1, SD = 18.0), as was the case in the original analysis in which speech rate was not controlled for.

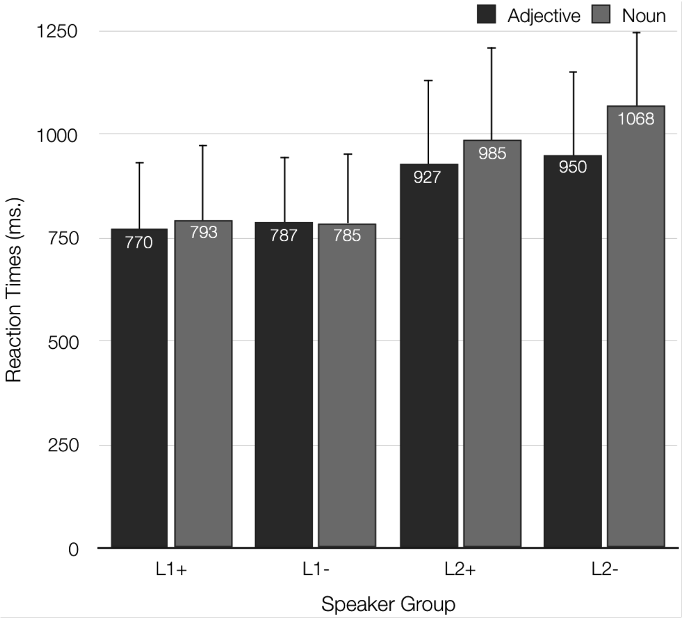


Figure 6. Mean RTs and standard deviations of L1 Dutch for utterances produced by L1 Dutch using matching pitch accent distributions, L1 Dutch using mismatching pitch accent distributions, proficient L2 Dutch, and less proficient L2 Dutch, separated by focus condition.

3.6.3 Discussion

The RT experiment shows that when examining absolute, raw RTs, L1 listeners respond slower to L2 speech than to L1 speech, which implies that speech by L2 speakers is less intelligible than L1 speech. This is as expected and corroborates both previous work using a RT paradigm (e.g., Munro & Derwing, 1995), as well as prior studies using a transcription paradigm (e.g., Derwing & Munro, 1997; Munro & Derwing, 1999). Furthermore, a distinction can be made based on the proficiency level of the L2 speaker: the higher the proficiency level of the speaker, the faster the speech can be processed by L1 listeners. This distinction in the perception of L1 and L2 speech in which pitch accent distributions are used to mark focus thus appears to mirror the distinction that has previously been reported in the production of these cues in L1 and L2 speech (Swerts & Zerbian, 2010; Chapter 2 of this dissertation).

However, contrary to our expectation, there was no significant difference between the raw RTs for L1 stimuli in which the pitch accent distribution corresponds to the original elicitation context and the RTs for those in which this is not the case. As the L1 stimuli in which the accent distributions reflect focus marking have the largest and most salient contrast between words that are marked by a pitch accent and words that are not, the fact that no difference is found between the two L1 conditions is intriguing. It suggests that even in the condition in which L1 listeners found it easiest to indicate which utterance was most natural in the preference task, intelligibility is not affected. This seems to suggest that when segmental factors are controlled for (i.e., none of the L1 speakers had a foreign accent), deviance in pitch accent distributions does not affect intelligibility.

Consequently, the difference between the RTs for L1+ or L1- stimuli on the one hand, and those for L2+ and L2- stimuli on the other hand, might be caused by other factors than deviance in pitch accent distributions, such as segmental deviance or speech rate. One might suggest that the substantial difference between the absolute RTs for stimuli by L2+ and L2- speakers could be explained solely by the increase of segmental deviance, which would corroborate earlier work by Munro and Derwing (1999), Terken and Lemeer (1988), and Vainio, Jarvikivi, Werner, Volk, & Valikangas (2002). When taking into account the accentedness ratings reported in the rating task, it is clear that L1 listeners rate less proficient L2 speech as substantially more accented than proficient L2 speech (it is a difference of 2.6 points, which is almost one third of the scale). However, the results of the analysis in which the RTs are measured as a proportion of the length of the stimulus utterance show that when speech rate is controlled for, all differences between language groups disappear. This implies that the differences in the RT that were initially observed between the languages group might also be due to the speech rate of the speakers rather than to their (in)correct use of pitch accent distributions to mark focus or the segmental quality of their speech. This is congruent with previous studies that have shown that speech rate affects accentedness and comprehensibility, and possibly also intelligibility (Kang, 2010; Munro & Derwing, 2001; Trofimovich & Baker, 2006).

Concluding, the current study is in line with previous work reporting that although prosodic deviance is easily recognized and judged, intelligibility may not be affected by it (Caspers & Horloza, 2012). The fact that our research does not reproduce earlier studies reporting the opposite (e.g., Anderson-Hsieh, Johnson & Koehler, 1992; Hahn, 2004; Magen, 1998; Swerts & Vroomen, 2015; Van Heuven et al., 1981), might be due to the fact that we used natural and semi-spontaneous speech stimuli where others relied on read or instructed speech, which is reported to have markedly different prosodic features and is characterized by a more pronounced difference between accented and deaccented words. In other words, while non-spontaneous speech stimuli may yield an effect of prosodic deviance on intelligibility measures, spontaneous speech stimuli may not necessarily do so. As oral communication is generally of a spontaneous nature, the importance of spontaneous speech stimuli in perception studies on prosodic and segmental deviance should not be ignored.

3.7 General discussion and conclusion

The current study investigated the effect of prosodic deviance in L1 and L2 semi-spontaneous speech on perceptions by L1 listeners with the aim of determining how deviance in pitch accent distributions used to mark focus affects perceived accentedness, comprehensibility, nativeness, and intelligibility. For each of these constructs an additional research question was whether the proficiency level of the speaker had a modulating effect. From the first study, it can be concluded that L1 listeners have unambiguous intuitions concerning the accentedness, comprehensibility, and nativeness of both L1 and L2 speech: They consistently rate L1 speech as less foreign accented, easier to comprehend and more nativelike than L2 speech. This effect is mirrored within L2 speech, where a similar tendency is observed between less proficient and proficient L2 speech. Interestingly, the data also reveal that comprehensibility is rated less severely than accentedness and nativeness are. While the difference between L1 speech ratings and less proficient L2 speech ratings is 6.4 and 6.6 points for accentedness and nativeness respectively, it is only 2.6 points for comprehensibility (with all ratings on the higher end of the scale), which suggests that even though L1 listeners make clear distinctions between the different speaker groups when it comes to sounding foreign, they feel the groups are less diverse regarding comprehensibility, which is in line with previous work by Munro and Derwing (1999) and Derwing and Munro (1997).

The notion that perceived accentedness and nativeness may be substantially more affected by deviance in pitch accent distributions than comprehensibility and intelligibility is further investigated in the last two experiments. These reveal that although L1 listeners are able to hear the difference between two utterances that only vary prosodically, and that they find utterances in which the prosody matches the focal structure significantly more natural than utterances in which focal structure and prosody do not match, this does not affect their intelligibility: L1 listeners do not process utterances without deviance in pitch accent distributions any faster than utterances with deviance in pitch accent distributions, when they are produced by L1 speakers. Although utterances produced by L2 speakers, irrespective of their proficiency level, yield significantly longer RTs than utterances produced by L1 speakers, this difference might be attributed to the fact that as L2 speech becomes less proficient, speech also is slower, which appears to hinder intelligibility. The results of both the analysis of the absolute RTs and the additional analysis in which speech rate is controlled for corroborate this assumption as the former shows that there is no difference between the RTs for L1+ and L1- stimuli (which have the same speech rate and only vary in pitch accent distribution), and the latter demonstrates that when speech rate is normalized, there are no differences in RTs between the either of the language groups, irrespective of their L1. These results are especially relevant when taking into account the stimuli that were used in the experiments. Earlier results on prosodic deviance often relied on read or instructed speech stimuli, which are known to have substantially different prosodic characteristics than spontaneous, natural speech stimuli

(Blaauw, 1994; Dellwo, et al., 2015; De Ruiter, 2015; Howell & Kadi-Hanifi, 1991; Laan, 1997; Swerts et al., 1996). Although many studies have reported significant effects of prosodic deviance either on accentedness, comprehensibility, nativeness or intelligibility in the past, very few of them used spontaneous speech stimuli, which may have biased their results. The fact that the results reported in these studies are not corroborated by the majority of the studies reviewed in the introduction appears to indicate that the distinction between non-spontaneous and spontaneous speech is an important one to keep in mind when conducting perception experiments.

Furthermore, our results may have pedagogical consequences as well: it has been suggested in previous work that prosodic cues deserve more attention in L2 classrooms (e.g., Mennen, 2007). Conversely, incorporating the acquisition of suprasegmentals in educational programs is often problematic as they are generally considered to be language features that are difficult to manipulate consciously: in part because in some cases they simply result automatically from a speaker's articulatory constraints (e.g., the declination effect as a natural consequence of decreasing air pressure in the lungs), and also because suprasegmentals belong to a class of features that tend to be minimally represented in writing systems, which makes it difficult for L2 learners to know when to produce them. Finally, the use of many prosodic cues, such as the distribution of pitch accents to mark focal structure, is also context-dependent, which might hinder successful acquisition. Taking this into account, the results of the present study are felicitous, as they suggest that at least some prosodic features might not be as relevant to successful communication in the L2 as previously expected. Depending on the objectives (e.g., reducing a foreign accent vs. successful communication), aptitude, and motivation of the L2 learner, this particular prosodic feature might be excluded from a didactic method in favour of cues that are more relevant to intelligibility for L1 listeners, such as acquiring fluency in the L2 (i.e., increasing the speech rate) or adequate pronunciation of certain L2 segments that may not be characteristic of the L1.

Naturally, our research could be extended in a number of ways. Although the current experiment made use of semi-spontaneous speech stimuli, for the reasons mentioned before, the use of this type of stimuli makes it difficult to control for possible confounds, such as item length and segmental deviance in the case of L2 speech. Although item length was controlled for in a later stage of the analysis, an ideal stimulus would vary in nothing but the prosodic cue under investigation. To our knowledge, the only way to produce such a stimulus would be to use synthetic speech stimuli, or spontaneous speech stimuli that have been manipulated synthetically to vary the pitch accent distributions. Doing this would involve the risk that speech would once again become less natural and spontaneous in nature, which would make it impossible to investigate accentedness, comprehensibility, nativeness, and intelligibility in one design, which was the objective of the present study. Thus, the inability to control for segmental factors appears to be inherent to the use of spontaneous speech stimuli. In future research, the current design might benefit from replication with synthetically adapted spontaneous speech stimuli.

Notes

[1] It is crucial to distinguish between two meanings of the word *accent*. First, one can refer to speech with

foreign characteristics as speech with a *foreign accent*, or as *accented*. Secondly, one can refer to an emphasized or stressed (element of a) word or sentence as *accented*, or as receiving a *pitch accent*. As these two terms overlap in choice of words, but are different in meaning, *pitch accent* will be used to refer to the latter, while *foreign accent(ed)*, or *accentedness* is used for the former.

2 Deaccentuation occurs when "a word that we might expect to be accented fails to be accented in a context where it has recently been used or where the entity to which it refers has recently been mentioned" (Ladd, 2006, p.175). In the example below one would expect the second mention of 'Mary' to receive a pitch accent according to general accentuation rules. However, 'Mary' is deaccented because she was mentioned in the preceding clause.

(1) John goes out with Mary, even though he knows I don't like Mary.

References

Akker, E., & Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism: Language and Cognition*, *6*, 81–96.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, *42*, 529–555.

Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning*, *38*, 561–613.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics, 32*, 233–250.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255-278.

Blaauw, E. (1994). The contribution of prosodic boundary markers to the perceptual difference between read and spontaneous speech. *Speech Communication, 14*, 359–375.

Boula de Mareüil, P., & Vieru-Dimulescu, B. (2006). The contribution of prosody to the perception of foreign accent. *Phonetica, 63*, 247–267.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. I. (1997). Training Japanese listeners to identify English /r/ and /l/: Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America, 101*, 2299–2310.

Caspers, J., & Horłoza, K. (2012). Intelligibility of non-natively produced Dutch words: Interaction between segmental and suprasegmental errors. *Phonetica*, *69*, 94–107.

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, United Kingdom: Cambridge University Press.

Cruttenden, A. (1993). The de-accenting and re-accenting of repeated lexical items. In *ESCA Workshop on Prosody* (pp. 16-19). Lund, Sweden.

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics*, *20*, 55-60.

Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*, 141–201.

Dellwo, V., Leemann, A., & Kolly, M. J. (2015). The recognition of read and spontaneous speech in local vernacular: The case of Zurich German. *Journal of Phonetics*, *48*, 13–28.

De Ruiter, L. E. (2015). Information status marking in spontaneous vs. read speech in story-telling tasks: Evidence from intonation analysis using GToBI. *Journal of Phonetics*, *48*, 29–44.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in Second Language Acquisition*, *19*, 1–16.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, *39,* 379–397.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, *48*, 393–410.

Doğruöz, A. S., & Backus, A. (2007). Postverbal elements in immigrant Turkish: Evidence of change? *International Journal of Bilingualism, 11*, 185–220.

Edmunds, P. (2009). *ESL speakers' production of English lexical stress: The effect of variation in acoustic correlates on perceived intelligibility and nativeness* (Doctoral dissertation). The University of New Mexico, New Mexico, United States.

E-prime (version 2.0.10.353). [Computer software]. (2015). Retrieved from http://www.pstnet.com/eprime.cfm

Face, T. L. (2000). Prosodic manifestations of focus in Spanish. *Southwest Journal of Linguistics, 19*, 45–62.

Face, T. L. (2002). *Intonational marking of contrastive focus in Madrid Spanish*. Munich, Germany: Lincom Academic Publishers.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly*, *39*, 399–423.

Flege, J. E. (1987). The production of new and similar phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*, 47–65.

Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication, 6*, 185–202.

Flege, J. E., Schirru, C., & MacKay, I. R. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication, 40*, 467–491.

Grover, C., Jamieson, D. G., & Dobrovolsky, M. B. (1987). Intonation in English, French and German: perception and production. *Language and Speech*, *30*, 277–295.

Gut, U., & Pillai, S. (2014). Prosodic marking of information structure by Malaysian speakers of English. *Studies in Second Language Acquisition, 36,* 283–302.

Gut, U., Pillai, S., & Mohd Don, Z. (2013). The prosodic marking of intonation status in Malaysian English. *World Englishes, 32*, 185–197.

Hagmann, L., & Dellwo, V. (2014). Listeners may rely on intonation to distinguish languages of different rhythm classes. *Loquens*, *1*, 1–11.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*, 201–223.

Heck, R. H., Thomas, S., & Tabata, L. (2013). *Multilevel modeling of categorical outcomes using IBM SPSS*. Oxford, United Kingdom: Routledge Academic.

Hoot, B. (2012). *Presentational focus in heritage and monolingual Spanish* (Doctoral dissertation). University of Illinois at Chicago, Chicago, United States.

Howell, P., & Kadi-Hanifi, K. (1991). Comparison of prosodic properties between read and spontaneous speech material. *Speech Communication, 10*, 163–169.

Hualde, J. I. (2005). *The sounds of Spanish*. Cambridge, Massachusetts, United States: Cambridge University Press.

IBM Corporation. (2015). SPSS Statistics (version 22) [Computer software]. Available from http://www.spss.com/software/statistics/

Jarvis, S. (2000). Semantic and conceptual transfer. *Bilingualism: Language and Cognition, 3*, 19–21.

Jiang, N. (2004). Semantic transfer and its implications for vocabulary teaching in a second language. *Modern Language Journal, 88*, 416–432.

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, *38,* 301–315.

Kang, O., Rubin, D. O. N., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, *94*, 554–566.

Kolly, M. J., & Dellwo, V. (2014). Cues to linguistic origin: The contribution of speech temporal information to foreign accent recognition. *Journal of Phonetics*, *42*, 12–23.

Krahmer, E., & Swerts, M. (2001). On the alleged existence of contrastive accents. *Speech Communication, 34*, 391–405.

Laan, G. (1997). The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a read speaking style. *Speech Communication, 22*, 43–65.

Ladd, D. R. (1996). *Intonational phonology*. Cambridge, Massachusetts, United States: Cambridge University Press.

LimeSurvey (version 1.92+). [Computer software]. (2014). Retrieved from https://www.limesurvey.org/nl/

Maassen, B., & Povel, D. J. (1984). The effect of correcting fundamental frequency on the intelligibility of deaf speech and its interaction with temporal aspects. *The Journal of the Acoustical Society of America, 76*, 1673–1681.

Magen, H. S. (1998). The perception of foreign-accented speech. *Journal of Phonetics*, *26*, 381–400.

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics, 32*, 543–563.

Mennen, I. (2007). Phonological and phonetic influences in non-native intonation. In J. Trouvain, & U. Gut (Eds.), *Non-native prosody: Phonetic description and teaching practice* (pp. 53-76). Berlin, Germany: Mouton de Gruyter.

Munro, M. J. (1995). Nonsegmental factors in foreign accent: Ratings of filtered speech. *Studies in Second Language Acquisition*, *17*(1), 17-34.

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, *38*, 289–306.

Munro, M. J., & Derwing, T. M. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning, 48*, 159–182.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *49*, 285–310.

Munro, M. J., & Derwing, T. M. (2001). Modelling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition, 23*, 451–468.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human perception and performance*, *24*, 756–766.

Nooteboom, S. G., & Kruyt, J. G. (1987). Accents, focus distribution, and the perceived distribution of given and new information: An experiment. *The Journal of the Acoustical Society of America, 82*, 1512–1524.

Pickering, L. (2001). The role of tone choice in improving ITA communication in the classroom. *TESOL Quarterly, 35,* 233–255.

Pickering, L. (2004). The structure and function of intonational paragraphs in native and nonnative speaker instructional discourse. *English for Specific Purposes*, *23*, 19–43.

Qualtrics. [Computer software]. (2014). Retrieved from http://www.qualtrics.com/

Ramírez Verdugo, M. D. (2002). Non-native interlanguage intonation systems: A study based on a computerized corpus of Spanish learners of English. *ICAME Journal, 26*, 115–132.

Ramus, F., Dupoux, E., & Mehler, J. (2003). The psychological reality of rhythm classes: Perceptual studies. In M. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 337–342). Barcelona, Spain: Universitat Autonomá de Barcelona

Rasier, L. (2006). Prosodie en vreemdetaalverwerving: accentdistributie in het Frans en in het Nederlands als vreemde taal ('Prosody and second language acquisition: accent distribution in French and Dutch as a foreign language', in Dutch) (Doctoral dissertation). Université catholique de Louvain, Louvain, Belgium.

Rasier, L., & Hiligsmann, P. (2009). Exploring the L1-L2 relationship in the L2 acquisition of prosody. *Online proceedings of first and second languages: Exploring the relationship in pedagogy-related contexts*, Oxford, United Kingdom. Retrieved from: http://www.education.ox.ca.uk/wordpress/wp-content/uploads/2010/08/RasierHiligsmann.doc.

Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research, 16*, 135–172.

Salamoura, A., & Williams, J. N. (2007). The representation of grammatical gender in the bilingual lexicon: Evidence from Greek and German. *Bilingualism: Language and Cognition, 10*, 257–275.

Schairer, K. E. (1992). Native speaker reaction to non-native speech. *The Modern Language Journal*, *76*, 309–319.

Swerts, M. (2007). Contrast and accent in Dutch and Romanian. *Journal of Phonetics*, *35*, 380–397.

Swerts, M., Krahmer, E., & Avesani, C. (2002). Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics*, *30*, 629–654.

Swerts, M., Strangert, E., & Heldner, M. (1996). F0 declination in read-aloud and spontaneous speech. In *Proceedings of the Fourth International Conference on Spoken Language*, *96*, 1501–1504.

Swerts, M., & Vroomen, J. (2015). Accent shifts in spoken noun phrases affect verification latencies of listeners in Dutch but not Canadian French. *Journal of Phonetics*, *52*, 170–182.

Swerts, M., & Zerbian, S. (2010). Intonational differences between L1 and L2 English in South Africa. *Phonetica*, *67*, 127–146.

Terken, J. (1984). The distribution of pitch accents in instructions as a function of discourse structure. *Language and Speech, 27*, 269–289.

Terken, J., & Lemeer, G. (1988). Effects of segmental quality and intonation on quality judgments for texts and utterances. *Journal of Phonetics*, *16*, 453–457.

Terken, J., & Nooteboom, S. G. (1987). Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes, 2*, 145–163.

Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28*, 1–30.

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, *15*, 905–916.

Ulbrich, C. (2013). German pitches in English: Production and perception of cross-varietal differences in L2. *Bilingualism: Language and Cognition, 16*, 397–419.

Vainio, M., Jarvikivi, J., Werner, S., Volk, N., & Valikangas, J. (2002). Effect of prosodic naturalness on segmental acceptability in synthetic speech. In *Proceedings of 2002 IEEE Workshop on Speech Synthesis*, 143–146.

Vallduví, E. (1991). *The informational component*. New York, New York, United States: Garland Publishing.

Van Donselaar, W., Koster, M., & Cutler, A. (2005). Exploring the role of lexical stress in lexical recognition. *The Quarterly Journal of Experimental Psychology, 58*, 251–273.

Van Els, T., & De Bot, K. (1987). The role of intonation in foreign accent. *The Modern Language Journal, 71*, 147–155.

Van Heuven, V. J. J. P., Kruyt, J. G., & De Vries, J. W. (1981). Buitenlandsheid en begrijpelijkheid in het Nederlands van buitenlandse arbeiders, een verkennende studie ('Non-nativeness and intelligibility of foreign workers in Dutch, an explorative study', in Dutch). *Forum der Letteren, 22*, 171–178.

Wennerstrom, A. (1994). Intonational meaning in English discourse: A study of non-native speakers. *Applied Linguistics*, *15*, 399–420.

Wennerstrom, A. (1998). Intonation as cohesion in academic discourse. *Studies in Second Language Acquisition*, *20,* 1–25.

Zampini, M. (2008). L2 speech production research: Findings, issues and advances. In J.G.H. Edwards & M.L. Zampini (eds.), *Phonology and second language acquisition* (pp. 219–249). Amsterdam, The Netherlands: John Benjamins Publishing.

Zubizarreta, M. L. (1998). *Prosody, focus, and word order.* Cambridge, Massachusetts, United States: MIT Press.

4

rhythm production

4

Learning direction matters

A study on L2 rhythm acquisition by Dutch learners of Spanish and Spanish learners of Dutch*

Abstract

This study examines the acquisition process of speech rhythm in Dutch learners of Spanish (DLS) and Spanish learners of Dutch (SLD) at different proficiency levels in order to determine whether learning direction affects the success of rhythm acquisition in a foreign language (L2). Analyses of lengthening effects showed that the two learner groups followed different developmental paths in their acquisition of accentual and final lengthening: both groups showed transfer effects from the L1, but while the DLS systematically approached their target until attainment, the SLD showed more variability in their development. In addition, syllable structure complexity affected L2 rhythm acquisition, and to a substantially larger extent for the SLD compared to the DLS. The results support a model of L2 rhythm acquisition in which learning direction is included as a factor, and that allows for the interaction of various language-specific properties that contribute to speech rhythm, like syllable structure complexity.

---

* **This chapter is based on:** Van Maastricht, L., Krahmer, E., Swerts, M., & Prieto, P. (2017). Learning direction matters: A study on L2 rhythm acquisition by Dutch learners of Spanish and Spanish learners of Dutch. Accepted for publication in *Studies in Second Language Acquisition.*

4.1 Introduction

The acquisition of a foreign language (L2) occurs within various linguistic dimensions simultaneously. While many studies focus on L2 attainment of segmental, lexical, syntactic, and semantic properties, research on the acquisition of L2 prosody is substantially underrepresented. Furthermore, within the field of L2 prosody acquisition, the attainment of L2 rhythm has received relatively little attention compared to other suprasegmental properties, like lexical stress, phrasal prominence, and speech rate (Gut, 2009). However, native listeners as young as five days old are capable of discriminating languages that have traditionally been classified as prototypically 'stress-timed' and 'syllable-timed', like Dutch and Spanish respectively, based only on their rhythm (Nazzi, Bertoncini & Mehler, 1998; Ramus & Mehler, 1999; Ramus, Dupoux & Mehler, 2003). Previous studies suggest that perceived rhythm is the result of an interaction between language-specific factors, such as timing properties, prominence and boundary marking by means of syllable duration, and syllable structure (e.g., Abercrombie, 1967; Li & Post, 2014; Prieto, Vanrell, Astruc, Payne & Post, 2012; Post & Payne, 2018; White & Mattys, 2007a). Thus, rhythm is related to phonemic, phonotactic, and intonational features of language, and producing speech with adequate rhythm therefore requires control in multiple areas.

Hence, it is perhaps unsurprising that many L2 learners have difficulties acquiring the speech rhythm characteristic of their L2. It has been shown that L2 speakers, especially in the early phases of learning, tend to transfer rhythmic properties of their native language (L1) to the L2 (e.g., White & Mattys, 2007b), suggesting that target-like rhythm production is easier when the L1 and L2 are rhythmically similar (e.g., Ordin & Polyanskaya, 2015). Indeed, the similarity between the L1 and L2 as a factor of successful L2 acquisition has been studied extensively within the fields of L2 phonology and phonetics, and several theoretical models are based on it, like the Second Language Model (SLM, Flege, 1995), the Perceptual Assimilation Model (PAM, Best, 1995), and the Second Language Perception model (L2LP, Escudero & Boersma, 2004). However, the *direction* in which languages are learned has been studied less frequently (Gut, 2009). Intuitively learning direction is an important factor to investigate: arguably acquisition is more challenging from less complex languages towards more complex ones, than vice versa. To our knowledge, the only study on learning direction as a function of L2 *prosody* acquisition is Rasier's (2006) study on the acquisition of pitch accents to mark focus by L1 Dutch learners of French and L1 French learners of Dutch, showing that learning direction indeed affected the degree of success with which L2 learners produced target-like pitch accent distributions. And while an analysis in which the same two languages are compared cross-directionally sheds more light on the processes underlying the role of learning direction in L2 acquisition, no study has performed such a comparison for speech *rhythm* acquisition.

Therefore, we explore whether the direction in which L2 acquisition occurs affects the successful attainment of speech rhythm by L2 learners of two languages that are rhythmically different, namely Dutch learners of Spanish (DLS) and Spanish learners of Dutch (SLD). In order to test which learner group advances more towards its target, we compare DLS and SLD with varying proficiency levels for two measures that correlate with speech rhythm, accentual and final lengthening, in different phonotactic conditions.

Before turning to our predictions, several concepts relevant to the current study are described in more detail.

4.2 Theoretical background

4.2.1 Speech rhythm

The construct of rhythm has been operationalized in various ways. A diachronic overview of rhythm analyses generally starts with the notion of rhythm as a categorical concept dependent on isochrony (Abercrombie, 1967; Pike, 1945). Within this view, a distinction has been made between syllable-timed and stress-timed languages, where the former refers to languages in which the intervals between the beginning of *all* syllables were taken to be equal (e.g., Spanish), while the latter applies to languages in which only the intervals between *stressed* syllables were assumed to be similar (e.g., Dutch). This categorical distinction was questioned by studies showing that the idea of equal intervals, between all syllables or between stressed syllables only, was not supported by acoustic measurements (e.g., Bolinger, 1965). This initiated a shift towards the notion of rhythm as a gradient property, with the underlying assumption that no language is either completely syllable- or stress-timed (Dauer, 1983,[1] 1987). The results of acoustic and phonetic experiments provided an overview of properties that are relevant to speech rhythm and enable comparisons between languages based on these properties.

More recently, it was shown that not only phonetic and phonotactic, but also prosodic features of language influence speech rhythm. Studies showed that languages differ in the extent to which they lengthen stressed and/or accented syllables vis-à-vis unstressed and unaccented syllables (e.g., Beckman & Edwards, 1994), as well as in their lengthening of syllables preceding a prosodic boundary within or at the end of an utterance (e.g., Byrd, 2000). Prieto et al. (2012) showed that the degree to which languages apply these prosodic lengthening measures contributes to cross-linguistic rhythmic differences. Resulting from these developments, more recent studies of speech rhythm rely on rhythm metrics to measure the timing patterns of utterances, occasionally in combination with lengthening analyses (Grabe & Low, 2002; Gut, 2009; Li & Post, 2014; Ordin & Polyanskaya, 2015; Prieto et al., 2012; Ramus, Nespor & Mehler, 1999). These analyses followed from the dismissal of rhythm as a dichotomous notion, leading to a need for quantitative data to corroborate the idea of a rhythm continuum and to position a given language on it.[2] In this study, we base our analyses of speech rhythm on measures of accentual and final lengthening, in agreement with Dauer's (1983, 1987) list of parametric criteria to rhythmically differentiate between languages, one of which is the presence or absence of durational variation between stressed and unstressed syllables and the use of pitch to mark prominence (also see Allen & Hawkins, 1978). The following section explains why the two languages studied in the current investigation differ significantly in their speech rhythm.

4.2.2 Typological differences between Dutch and Spanish

Several typological differences between Spanish and Dutch have been hypothesized to underlie the perceptual distinction between these languages concerning rhythm. One differences concerns syllable complexity constraints: the majority of Spanish syllables have an open structure (syllables consisting of a consonant (C) followed by a vowel (V) are most frequent in two corpus studies: 58.0% of all syllables had CV structure in Navarro Tomás, 1966, and 53.9% in Hartsuiker, 2002), while the majority of Dutch syllables is closed (CVC syllables represented 62.4% in the corpus study by Hartsuiker, 2002). Moreover, Spanish allows for relatively few syllable structures that are more complex than the CV configuration. Navarro Tomás (1966) stated that the most complex syllable type in Spanish is CCVCC, as found in the first syllable of *trans-for-mar*, 'to transform'. Conversely, Dutch is documented as more varied in its syllable structure with complex structures being the norm. Syllable complexity can increase to up to 7 segments in one syllable, for instance in the word *strengst* ('strictest'), which has a CCCVCCC structure (Booij, 1995; Van Zon, 1997). Since Dutch and Spanish differ typologically in this respect, the current study controls for syllable structure in two out of the three conditions (using predominantly CV and CVC syllables respectively), while in the last (Mixed) condition syllable structures are used that are typical of both languages. In addition to these phonotactic differences, the two languages also differ in prosodic properties: Spanish is known to employ little accentual and final lengthening while both are employed extensively in Dutch (Cambier-Langeveld, Nespor & Van Heuven, 1997; Cambier-Langeveld & Turk, 1999; Delattre, 1966; Prieto et al., 2012).[3] In the following section, the effects of these differences on L2 rhythm acquisition are discussed.

4.2.3 L2 rhythm acquisition

Prior work on L2 rhythm attainment generally concentrated on the influence of the L1 on the L2, and typically reported that although L2 learners from different L1 backgrounds increasingly approached their target, considerable transfer effects usually also occur from the L1 to the L2 (e.g., Carter, 2005; White & Mattys, 2007a). Recently, Li and Post (2014) investigated the rhythm produced by Chinese and German learners of English with intermediate or advanced proficiency level. Their analyses showed that while learners from both L1 backgrounds produced rhythm metric values and syllable durations that increasingly approached the L2 target, their development also showed signs of L1 transfer: where intermediate learners produced values that were closer to those typical of their L1, the advanced learners produced values that were more similar to those of the L2 target. Interestingly, both learner groups performed equally well, which is surprising, because intuitively German rhythm is more similar to English than to Mandarin rhythm. One might therefore assume that the German learners of English would be more successful at producing the target speech rhythm than the Chinese learners of English.

This is precisely the idea developed further in Ordin and Polyanskaya (2015), who compared French and German L2 learners of English at beginner, intermediate, or advanced proficiency level. Their results corroborated those of Li and Post (2014) in that rhythm metric values of both learner groups revealed that durational variability increased as L2 acquisition progressed, which would be an indication of universal L2 acquisition development. Conversely, their results further showed that while the most proficient German learners of English achieved target values (and for some metrics the intermediate learners did too), the French learners of English did not. Ordin and Polyanskaya considered this an indication that L1 speakers of a syllable-timed language (here French) encountered more difficulty acquiring the speech rhythm of a stress-timed language (here English), than L1 speakers of another stress-timed language (here German). However, since Ordin and Polyanskaya compared two different L1-L2 combinations, the design of their study makes it impossible to rule out the possibility that the differences between these two learner groups were due to other segmental, phonotactic or prosodic properties in which French and German differ from each other.

4.3 The current study

In view of our limited understanding of rhythm transfer in general, and the importance of learning direction in this context specifically, this study compares DLS and SLD in their rhythm production, to determine which L2 group is more successful at producing target-like speech rhythm. Consequently, a language model is required that allows for predictions based on more than just the similarity of the L1 and L2, since both learning directions consist of the same language combination. Unfortunately, this excludes popular models of L2 acquisition, such as the SLM (Flege, 1995), PAM (Best, 1995), and L2LP (Escudero & Boersma, 2004). Moreover, these models concern the acquisition of segmental features and are therefore difficult to apply to suprasegmental L2 properties. Other models that do allow for predictions regarding prosodic features tend to focus on prosodic cues at a lexical level only, and generally take a Universal Grammar perspective, assuming that specific parameters are organized into a hierarchical tree structure in which some are embedded within others (e.g., Archibald, 1994; Özçelik, 2016).

We therefore base our predictions on Eckman's (1977; 2008) Markedness Differential Hypothesis (MDH), which is applicable to most areas of L2 acquisition, and does not depart from a specific language acquisition theory:

(1) MDH: "The areas of difficulty that a language learner will have can be predicted such that
   (a) Those areas of the target language which differ from the native language and are more marked that the native language will be difficult;
   (b) The relative degree of difficulty of the areas of difference of target language which are more marked that the native language will correspond to the relative degree of markedness;
   (c) Those areas of the target language which are different from the native language, but are not more marked than the native language will not be difficult."

(Eckman, 1977, p. 321)

Eckman defined markedness as follows: "A phenomenon is more typologically marked if its presence in a language implies the presence of another phenomenon; but the presence of the latter does not imply the presence of the former" (1977, p. 320-321).

As argued, Dutch and Spanish not only differ concerning the overall perception of their rhythm, but also with respect to various phonotactic and prosodic properties that underlie this perceptual distinction. The MDH can therefore be applied on at least three levels: First, young children initially produce speech with a rhythm that has been classified as more syllable-timed, and only later acquire the rhythmic properties specific of their L1 (Allen & Hawkins, 1978; Bunta & Ingram, 2007; Grabe, Watson & Post, 1999; Schmidt & Post, 2015). Most recently, Polyanskaya and Ordin (2015), investigated the attainment of rhythmic patterns by monolingual English children from 4-5 to 10-11 years old and adults. Their results corroborated earlier work and showed that the speech rhythm of the children developed from more syllable-timed to more stress-timed as language acquisition continued. As we know of no cases where infants first produced a stress-timed rhythm (to later develop a syllable-timed speech rhythm if this is typical of their L1), we assume that a stress-timed rhythm implies a syllable-timed rhythm in an earlier developmental stage, but not vice versa, indicating that stress-timed rhythm is typologically more marked than syllable-timed rhythm.

Second, a similar reasoning is applicable to correlates of rhythm, such as syllable complexity (Prieto et al., 2012): the use of complex syllable structures such as CCCVCCC implies that simple syllable structures, such as CV, are also possible within a language (an example from Dutch being *da-me*, 'lady'). However, the possibility of a CV syllable in Spanish does not imply that a syllable with CCCVCCC structure is also acceptable. From this it follows that the syllable structure of Dutch is more marked than the syllable structure of Spanish (also see Levelt & Van de Vijver, 2004; Ordin & Polyanskaya, 2015). Third, Dutch is also more marked than Spanish concerning lengthening effects, which are known to correlate with rhythm perception, as accentual and final lengthening are employed more extensively in Dutch than in Spanish. Lengthening implies a baseline that is not lengthened, but not vice versa: Not only does lengthening require more physiological effort than not lengthening (Ten Bosch, 1991), but the majority of all syllables in speech is not lengthened, whereas only a subset of the syllables is lengthened. This implies that the former is indeed the 'norm' (less marked), while the latter is the 'exception' (more marked). In sum, in all areas discussed, Dutch is arguably more marked than Spanish, which, according to the MDH, should make acquisition of these properties more difficult in Dutch than in Spanish. We therefore predict the following:

(2) H: Dutch learners of Spanish (DLS) are more successful at approaching their target rhythm than Spanish learners of Dutch (SLD).

Recently, the MDH has been used in two studies on L2 prosody acquisition: Rasier (2006), who applied it to the production of (de)accentuation patterns to signal focus in L2 French by L1 speakers of Dutch and L2 Dutch by L1 speakers of French, and Ordin and Polyanskaya (2015), who employed it in their analysis of L2 rhythm acquisition by German and French learners of English. Both reported that learners with an L1 background that is less marked than the target language (the L1 speakers of French who were learning Dutch, and the French learners of English, respectively) were less successful at attaining the L2 target than learners with an L1 background that is more or equally marked as the target L2 (the Dutch learners of French, and the German learners of English, respectively). In the next section, the collection and analysis of speech data by DLS, SLD and L1 speakers of Spanish and Dutch is described, followed by a comparison of the two learner groups, by means of accentual and final lengthening measures.

4.4 Method

4.4.1 Participants

70 adults participated in our experiment: 5 L1 speakers of Dutch and 5 L1 speakers of Spanish, whose

data serve as a baseline[4] to which the data of 30 DLS and 30 SLD are compared. All participants were raised in a monolingual environment and participated voluntarily (**Table A1** in the appendix contains those details about the speaker sample that are relevant to the experiment). The DLS were students of the Spanish program at the University of Groningen or Fuentes Academia de Español. The most proficient DLS were teachers at the Spanish Department of the Radboud University in Nijmegen or the University of Groningen. The SLD were students at the Escuela Oficial de Idiomas in Madrid or Barcelona, and the most proficient SLD were generally teachers at the Escuela Oficial de Idiomas. The L2 learners were subdivided into different proficiency groups, based on the proficiency levels of the Common European Framework of Reference for Languages, which distinguishes between six different proficiency levels ranging from A1 or A2 for beginners, to B1 and B2 for intermediate learners, and C1 and C2 for advanced speakers of an L2 (Council of Europe, 2001). Five speakers were recorded per proficiency level. The aforementioned institutions already used these proficiency levels, which facilitated the process of determining the proficiency of our participants. Their level in this study corresponded to the level of the last course they had successfully completed. Participants were asked to self-evaluate their skills with respect to specific reading, writing, speaking and listening proficiency, which were corroborated by the first author with their teachers. In general, these were congruent with students' overall level, with the productive skills being slightly more challenging than the receptive skills.

Since French is an obligatory subject in Dutch high schools, as is English in Spanish high schools, all participants had some knowledge of an additional West Germanic or Romance language. However, none of them spoke that language at a proficiency level higher than their target L2 proficiency level. We therefore assume that L2 learners were not influenced in their target rhythm production by other foreign languages from the same language family.

4.4.2 Materials

Following Prieto et al. (2012), the stimuli consisted of 30 sentences per language: 5 sentences with predominantly open syllables (CV), 5 with mostly closed syllables (CVC), and 20 that reflected typical syllable structures in either Dutch or Spanish (Mixed). Consequently, syllable structure was controlled in one third of the stimuli. The Spanish CV and CVC sentences were taken from Prieto et al. (2012). The Dutch CV and CVC sentences were created by the authors to match the Spanish ones. The mixed sentences were taken or adapted from Nazzi, et al. (1998) and Prieto et al. (2012). The percentage of open syllables was 81.6% in the Dutch CV sentences and 91.9% in the Spanish CV sentences. In the Dutch CVC sentences 78.3% of the syllables were closed, while in the Spanish CVC sentences 59.0% were closed. In the mixed sentences, 47.7% of the Dutch sentences had an open structure, in contrast to the Spanish mixed sentences with 69.1% of all syllables open. Thus, the manipulation of syllable structure was realized as intended.[5] All sentences were matched for number of syllables (range: 12-19 syllables), although this may vary somewhat across individuals as a result of participant-specific pronunciation preferences. Sentences were also matched as best as possible for orthographic words (Spanish M=9.03, Dutch M=9.63 per sentence) and prosodic words (Spanish M=4.87, Dutch M=5.26 per sentence).

Infrequent words and complex sentence constructions were avoided where possible, to facilitate the task for L2 learners. Example (3) shows a stimulus sentence for each of the categories in Dutch and Spanish. The whole stimuli set can be found in the appendix.

(3a) CV syllable structure (16 syllables):
Dutch: De mama van Susana is een gezellige lerares
Spanish: La madre de Susana es una buena profesora

(3b) CVC syllable structure (15 syllables):
Dutch: De wedstrijd van de voetbalclub was niet in het sportcomplex
Spanish: El mitin del club de tenis no fue en el parking del club

(3c) Mixed syllable structure (16 syllables):
Dutch: De dader werd helaas bij gebrek aan bewijs vrijgesproken
Spanish: Reportan inundaciones graves en la primavera

4.4.3 Procedure

Experimental sessions were performed individually and lasted approximately 10 minutes for the L1 speakers, and 20 minutes for the DLS and SLD, who performed the task in both their L1 and L2. The order in which the L2 learners performed these tasks was randomized across participants. The recordings, made with Praat (Boersma & Weenink, 2015) and the internal microphone of an Apple Macbook Pro, took place in a quiet room. Participants were instructed to read the sentences at a normal, comfortable pace from the laptop screen, and to repeat the sentence if there were hesitations or other irregularities in their speech, continuing to the next sentence at their own convenience. While a higher L2 proficiency level generally entailed less repetitions, this method ensured very few disfluencies in the speech by L2 learners of all proficiency levels. The few pauses and/or disfluencies that were unavoidable in the recording of the data, were excluded from measurement on syllable basis and are not included in data analysis. L2 learners could ask for translations of words and sentences, but the experiment leader did not provide phonetic coaching, and refrained from pronouncing the target words herself. Participants filled in a questionnaire to verify that they met the requirements of each language group concerning L1/L2, proficiency, experience in countries where the target L2 is spoken, age and gender, and to ensure that none of the participants had dyslexia or visual problems, which might influence their reading performance.

4.4.4 Prosodic analysis

The audio recordings were analyzed prosodically in Praat: each utterance was segmented into words, syllables, and phonemes. Segmental annotation for all utterances was first performed automatically using Praatalign, version 1.9b (Lubbers & Torreira, 2015). Subsequent segmentation and coding was performed manually by the first author, a trained phonetician who is an L1 speaker of Dutch and proficient in Spanish. Manual correction of the preprocessed speech was done by visual inspection of the speech waveforms and wideband spectrograms following standard criteria (see Peterson & Lehiste, 1960; White & Mattys, 2007a; Prieto, et al., 2012).

In two additional tiers, segments were coded as consonants or vowels, and syllable boundaries were placed. For Spanish, these boundaries were positioned following Prieto et al. (2012): prevocalic glides were coded as part of the preceding consonantal interval, and postvocalic glides as part of the preceding vocalic interval (e.g., the first syllable of *buena* was treated as CCV; the first syllable of *Ceilán* as CVV). Furthermore, CV structures were maintained whenever possible and a CV resyllabification process occurred across word boundaries. Following Schiller, Meyer, Baayen, and Levelt (1996), resyllabification also took place in the Dutch utterances, after taking into account phonological rules such as final devoicing (‘*aard*’ /ard/ becomes [art]), degemination (‘*komen naar*’ /kɔmən nar/ becomes [kɔmənar]), as well as final –n deletion after a schwa, and progressive voice assimilation (‘*uitvallen*’ /œytvɑlən/ becomes [œytfɑlə]).

In order to analyze final lengthening effects, each syllable was marked for its phrasal position as either non-final, intermediate phrase (ip) final or intonational phrase (IP) final following the procedure described in Prieto et al. (2012). The criterion for an IP break was a pause of at least 200 milliseconds, while a break of less than 200 milliseconds and a continuation rise characterized an ip boundary. The non-final syllables were then taken as a baseline condition to which the length of ip-final and IP-final syllables were compared. Prosodic prominence was also annotated, distinguishing between unstressed and unaccented, stressed and accented, and stressed and nuclear accented syllables. In this case, unstressed and unaccented syllables correspond to the baseline to which stressed and accented, and stressed and nuclear accented syllables were compared.[6] On the next page, **Figure 1** (on the next page) illustrates the orthographic, segmental, and prosodic transcription of the Spanish utterance *La madre de Susana es una buena profesora* (‘Susana’s mother is a good teacher’), produced by an L1 speaker of Spanish. The first tier contains the orthographic transcription, the second one the phonetic segmentation, and the third the consonant/vowel coding. In the fourth tier, syllabic segmentation and syllable structure is depicted, and in the two final tiers prominence and phrasal position is coded.[7] In total, 2.100 utterances were collected (5 speakers × 30 utterances × 14 language groups), resulting in 35.808 analyzed syllables, and 48.068 analyzed segments.

Inter-transcriber reliability of the prosodic coding was tested with ten percent (105 Dutch and 105 Spanish utterances) of our data. These utterances were randomly selected by the first author, who ensured that they equally represented all language groups, speakers, and phonotactic conditions. After discussing several examples with the first author, two transcribers (one L1 speaker of Dutch, and one L1 speaker of Spanish) independently labeled the utterances for phrasal position and phrasal prominence using the guidelines provided in this section. A comparison of the prosodic transcription across the two transcribers per language revealed a high inter-rater reliability both in phrasal prominence and phrasal position labeling. Agreement on the choice of phrasing level was high: 99.1% consistency for Dutch ($\varkappa$ = .974), and 93.4% for Spanish ($\varkappa$ = .785). Similarly, agreement on the choice of phrasal prominence levels was 97.8% for Dutch ($\varkappa$ = .956), and 88.5% for Spanish ($\varkappa$ = .754). This is comparable to inter-rater reliability scores in similar studies using prosodic labeling (Prieto et al, 2012), indicating that both prosodic features were labeled reliably (Landis & Koch, 1977).

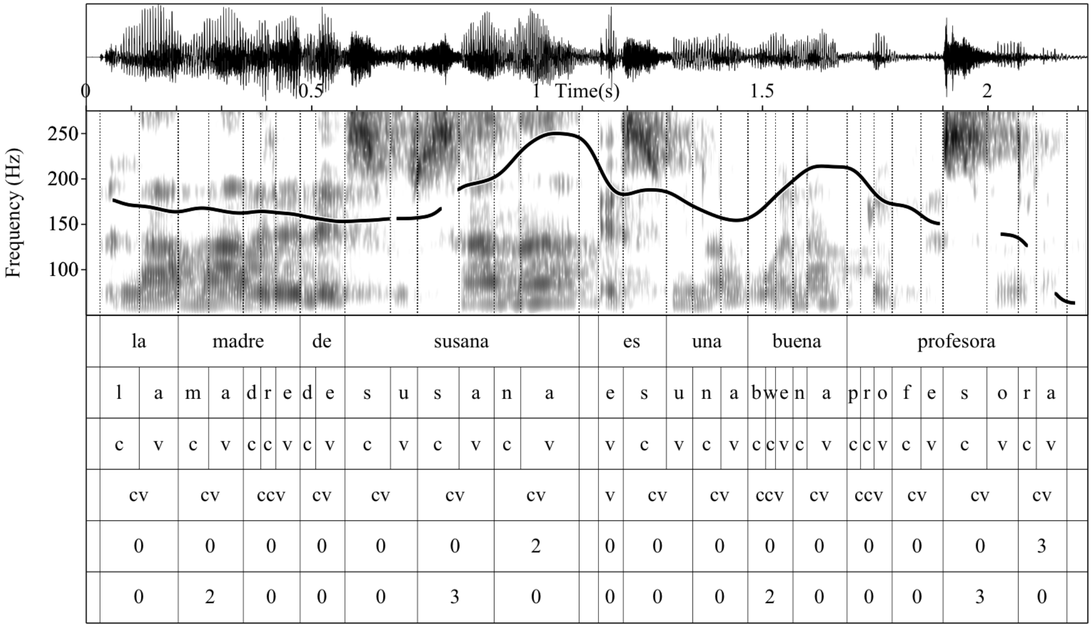


Figure 1. Waveform, spectrogram, F0 contour, and labeling scheme used for the Spanish utterance *La madre de Susana es una buena profesora,* 'Susana's mother is a good teacher'.

4.5 Results

In what follows, we first compare syllable duration data by L1 Dutch and L1 Spanish as a function of prosodic prominence and phrasal position, to form a baseline against which we subsequently compare the DLS and SLD. All analyses are performed in IBM SPSS Statistics 22 (IBM, 2015) using a Generalized Linear Mixed Model (GLMM). Specific response variables and fixed factors are described in the relevant sections, but for all analyses subjects and items were included as random factors, including random intercepts, and random slopes for fixed effects and their interaction (Barr, Levy, Scheepers, and Tily, 2013). Pairwise comparisons that explain main effects and interactions were Bonferroni adjusted.

4.5.1 L1 Spanish versus L1 Dutch

A GLMM analysis was performed with syllable duration in seconds as the response variable, and Language Group (2 levels: L1 Dutch, L1 Spanish), Syllable Structure (3 levels: CV, CVC, Mixed), Phrasal Prominence (3 levels: unstressed and unaccented, stressed and accented, stressed and nuclear accented), and Phrasal

Position (3 levels: non-final, ip-final, IP-final) as fixed factors. The analysis reveals significant main effects for all fixed factors and significant interactions for all relevant combinations, except the interaction between Language Group and Phrasal Position (see **Appendix Table A2** for all potential main effects and interactions). Pairwise comparisons between the three Phrasal Prominence conditions within each L1 group reveal that in both L1 Spanish and L1 Dutch increasing prominence of the syllable entails longer syllable durations. In L1 Spanish, all Phrasal Prominence levels differ significantly from one another ($p < .001$), whereas in L1 Dutch, both stressed and accented syllables, and stressed and nuclear accented syllables are significantly longer than unstressed and unaccented syllables ($p < .001$), but the syllable durations of stressed and accented syllables do not differ significantly from those of nuclear accented syllables ($p = .099$), see **Figure 2.**

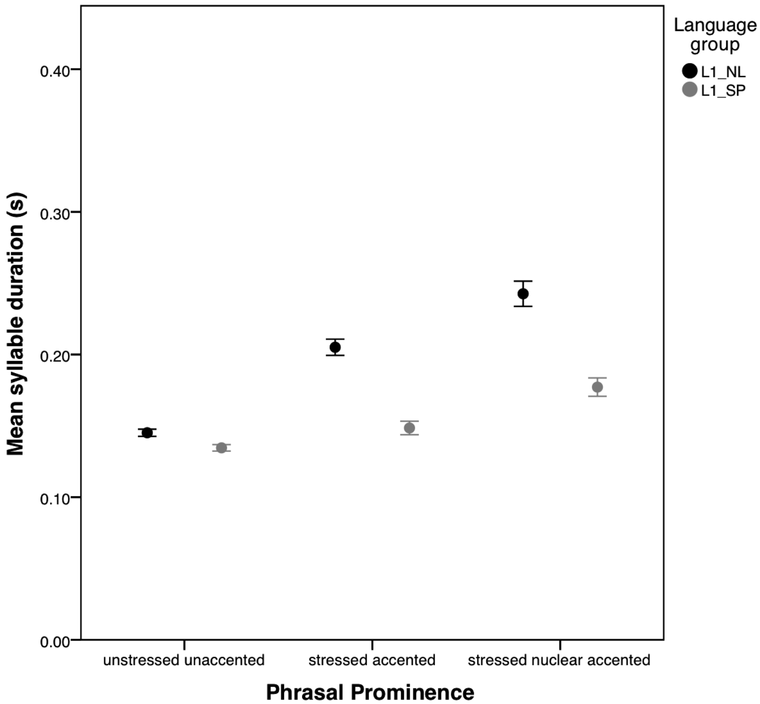


Figure 2. Mean syllable duration (in milliseconds) in L1 Spanish and L1 Dutch, separated by Phrasal Prominence condition. All sentences.

Pairwise comparisons between Language Groups within Phrasal Prominence levels reveal that L1 Dutch and L1 Spanish have similar default syllable lengths for unstressed and unaccented syllables (p = .652), but they differ significantly from each other for the other two Phrasal Prominence levels (p < .001) as syllables are lengthened more extensively in L1 Dutch than in L1 Spanish. This confirms prior research on the degree of accentual lengthening used in both languages (Cambier-Langeveld, et al., 1997; Cambier-Langeveld & Turk, 1999; Delattre, 1966; Prieto et al., 2012).

Controlling for Syllable Structure by examining the CV sentences only does not generate substantial differences to this pattern, see **Appendix Table A3** for mean syllable durations per Phrasal Prominence condition and Language Group for both CV and all sentences. The only difference is that the values are lower for both Language Groups in the CV condition than in the complete dataset, which can be explained by the fact that in CVC and Mixed sentences syllables are usually longer, due to their more complex syllable structure. Pairwise comparisons between Language Groups within each Phrasal Prominence level for CV sentences only reveal a similar pattern as the one found for all sentences: L1 Dutch and L1 Spanish have comparable syllable lengths for unstressed and unaccented syllables (p = .205), but they differ significantly from each other for the other two Phrasal Prominence levels (stressed and accented syllables: p = .003, stressed and nuclear accented syllables: p = .009), see **Figure 3**.

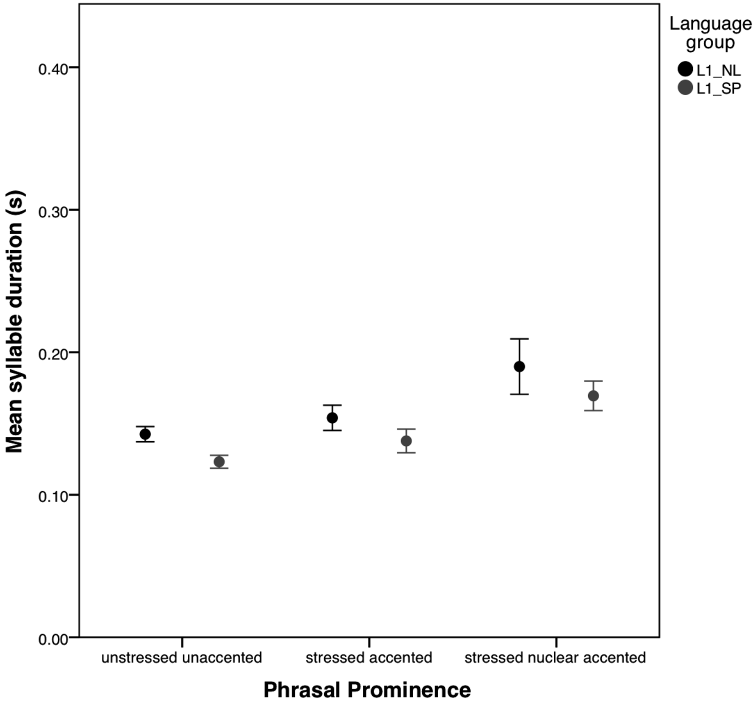


Figure 3. Mean syllable duration (in milliseconds) in L1 Spanish and L1 Dutch, separated by Phrasal Prominence condition. CV sentences only.

Regarding final lengthening, pairwise comparisons between the three Phrasal Position conditions within L1 groups show that for both Language Groups syllable durations increase significantly with respect to the baseline when the phrasal position of a syllable precedes an ip or IP boundary, see **Figure 4** on the next page ($p < .001$ for all comparisons). Furthermore, in both Language Groups the ip-final and IP-final syllables do not differ significantly from each other (Dutch: $p = .863$, Spanish: $p = .374$). Pairwise comparisons between Language Groups within each Phrasal Position condition show that L1 Dutch and L1 Spanish differ significantly from each other for all Phrasal Position conditions (non-final and IP-final syllables: $p < .001$, ip-final syllables: $p = .003$). This could again be due to the fact that syllables are longer in L1 Dutch in general, even in non-final position, due to its more complex syllable structure.

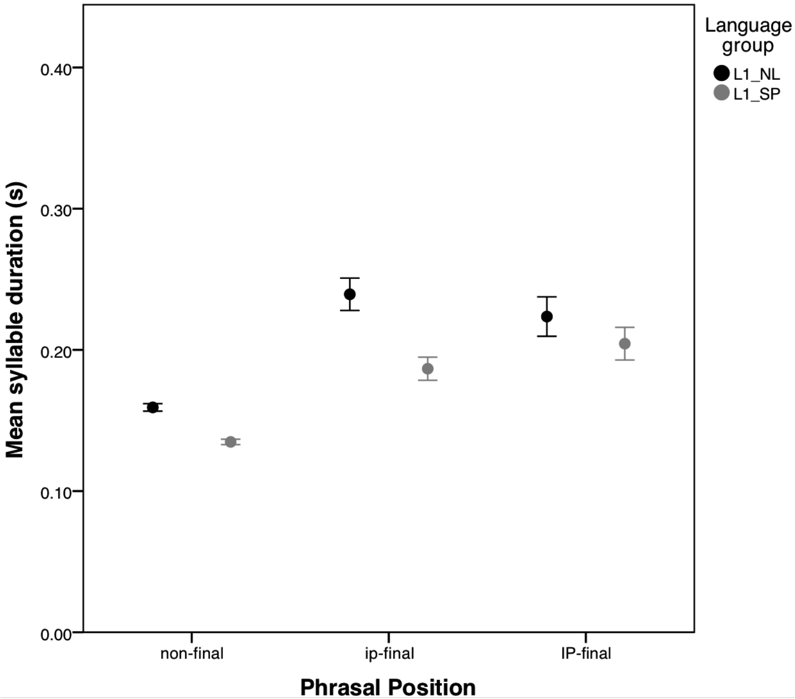


Figure 4. Mean syllable duration (in milliseconds) in L1 Spanish and L1 Dutch, separated by Phrasal Position condition. All sentences.

Controlling for Syllable Structure by examining pairwise comparisons between Phrasal Position conditions within both Language Groups for the CV sentences only, reveals a comparable pattern: in both L1s, non-final syllables are significantly shorter than ip-final and IP-final syllables ($p < .001$ for both), while there is no significant difference between ip-final and IP-final syllables (Dutch: $p = .908$, Spanish: $p = .434$). Pairwise comparisons show that speakers of L1 Dutch and L1 Spanish still differ significantly from each other in the non-final and IP-final conditions (non-final syllables: $p = .007$, IP-final syllables: $p = .021$), but the difference between the two L1s is not significant in the ip-final condition ($p = .313$), see **Figure 5**. **Appendix Table A4** contains the mean syllable durations per Phrasal Position and Language Group for all sentences and CV sentences only.

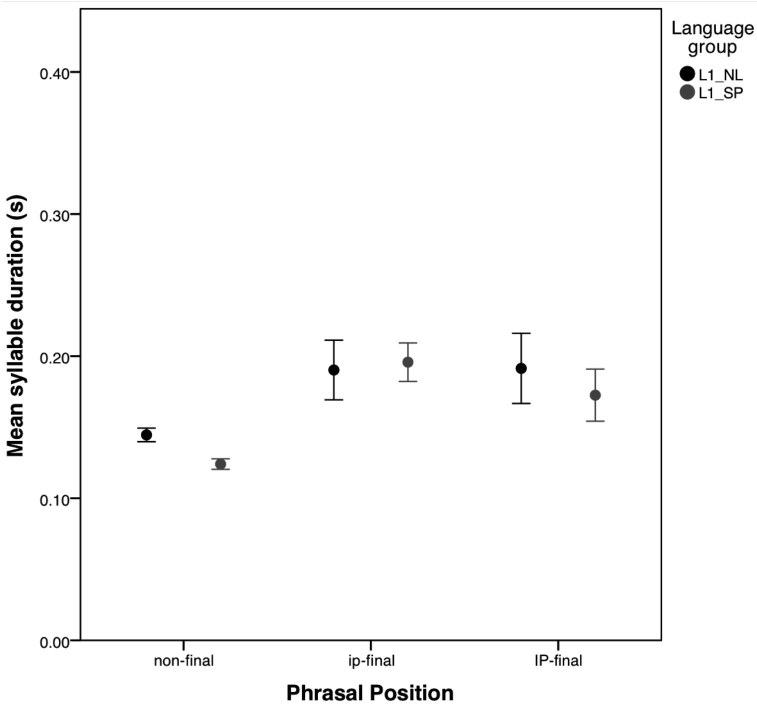


Figure 5. Mean syllable duration (in milliseconds) in L1 Spanish and L1 Dutch, separated by Phrasal Position condition. CV sentences only.

The significant interaction effect between Phrasal Position and Phrasal Prominence on syllable durations was further explored by examining the mean syllable durations per Language Group for all Phrasal Position and Phrasal Prominence combinations. **Table 1** shows that both factors interact systematically: within each Phrasal Position condition accentual lengthening effects increase as syllables are more prominent in the sentences, while increasing syllable durations are also observed between Phrasal Position conditions.

Table 1 *Mean syllable durations in seconds (standard error) in seconds for all sentences produced by speakers of L1 Dutch and L1 Spanish, separated per Phrasal Position and Phrasal Prominence combination (*N *= 10)*

Language Group	Phrasal Position	Phrasal Prominence	*M* syllable duration (*SE*)
L1 Dutch (*N* = 5)	non-final	unstressed, unaccented	.14 (.05)
		stressed, accented	.20 (.06)
		stressed, nuclear accented	.21 (.07)
	ip-final	unstressed, unaccented	.21 (.07)
		stressed, nuclear accented	.29 (.07)
	IP-final	unstressed, unaccented	.19 (.06)
		stressed, nuclear accented	.31 (.09)
L1 Spanish (*N* = 5)	non-final	unstressed, unaccented	.13 (.04)
		stressed, accented	.15 (.05)
		stressed, nuclear accented	.16 (.05)
	ip-final	unstressed, unaccented	.18 (.05)
		stressed, nuclear accented	.21 (.04)
	IP-final	unstressed, unaccented	.19 (.07)
		stressed, nuclear accented	.24 (.05)

4.5.2 L1 Spanish versus DLS

To compare the DLS to their target L1 group, a GLMM analysis was performed with syllable duration as the response variable, and Language Group (7 levels: L1 Spanish, DLS_A1, DLS_A2, DLS_B1, DLS_B2, DLS_C1, DLS_C2), Syllable Structure (as above), Phrasal Prominence (as above), and Phrasal Position (as above) as fixed factors. The analysis reveals significant main effects for all fixed factors and significant interactions for all relevant combinations, except for the interaction between Language Group, Syllable Structure, and Phrasal Position, (see **Appendix Table A5** for all potential main effects and interactions).

Pairwise comparisons between all Language Groups overall reveal that the DLS gradually approach target syllable durations as their proficiency increases. The most proficient group, DLS_C2, no longer differs significantly from the target L1 Spanish (p = .735), while all other DLS groups still do (DLS_C1 and DLS_B2: p = .001, DLS_B1, DLS_A2, and DLS_A1: p

< .001). This implies that while the DLS_C2 learners have attained a native-like level in their L2, learners of all other levels still differ significantly from their target. Controlling for syllable structure by comparing the DLS to the L1 Spanish within the CV condition reveals that the effect of Language Group is partially dependent on syllable structure: within the CV condition the L1 Spanish values are not only significantly different from those of the DLS_C2 (p = 1.000), but also from those of the DLS_C1 (p = .116) and DLS_B2 (p = .064). To examine whether the DLS approach L1 values similarly for accentual and final lengthening, pairwise comparisons between Language Groups within prominence and finality conditions were performed.

Regarding Phrasal Prominence, the results show that within all Phrasal Prominence conditions the DLS_C2 are not significantly different from the L1 Spanish, see **Table 2** and **Figure 6**.

4

Table 2 p-*values of pairwise comparisons between the L1 Spanish (*N *= 5) and all DLS groups (*N *= 30) for Phrasal Prominence, separated by syllable structure*

Pairwise comparison	All sentences			CV sentences only		
	Unstressed, unaccented	Stressed, accented	Stressed, nuclear accented	Unstressed, unaccented	Stressed, accented	Stressed, nuclear accented
L1 Spanish – DLS_A1	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p < .001$
L1 Spanish – DLS_A2	$p < .001$	$p < .001$	$p < .001$	$p = .013$	$p = .001$	$p < .001$
L1 Spanish – DLS_B1	$p < .001$	$p < .001$	$p < .001$	$p = .033$	$p < .001$	$p < .001$
L1 Spanish – DLS_B2	$p = .014$	$p = .001$	$p < .001$	$p = .337$	$p = .017$	$p = .164$
L1 Spanish – DLS_C1	$p = .004$	$p = .004$	$p < .001$	$p = .448$	$p = .064$	$p = .299$
L1 Spanish – DLS_C2	$p = 1.000$	$p = .964$	$p = .326$	$p = 1.000$	$p = 1.000$	$p = 1.000$

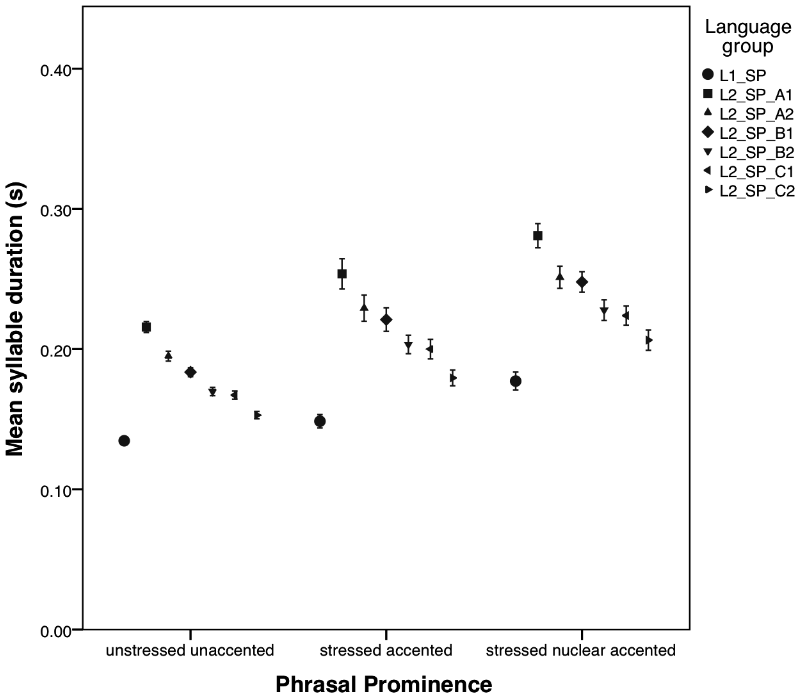


Figure 6. Mean syllable duration (in milliseconds) in L1 Spanish and DLS of all proficiency levels, separated by Phrasal Prominence condition. All sentences.

Contrary to the L1 data, this effect appears susceptible to the syllable structure of the sentence in speech by L2 learners, as making the same comparisons within the CV condition reveals that the three highest proficiency levels are comparable to the L1 target, see **Table 2** and **Figure 7**. Examination of the syllable durations of the different Phrasal Prominence conditions within all Language Groups reveals that all DLS groups show a similar pattern to the L1 Spanish, in which syllable durations are longer as syllables are more prominent within an utterance, see **Appendix Table A6**.

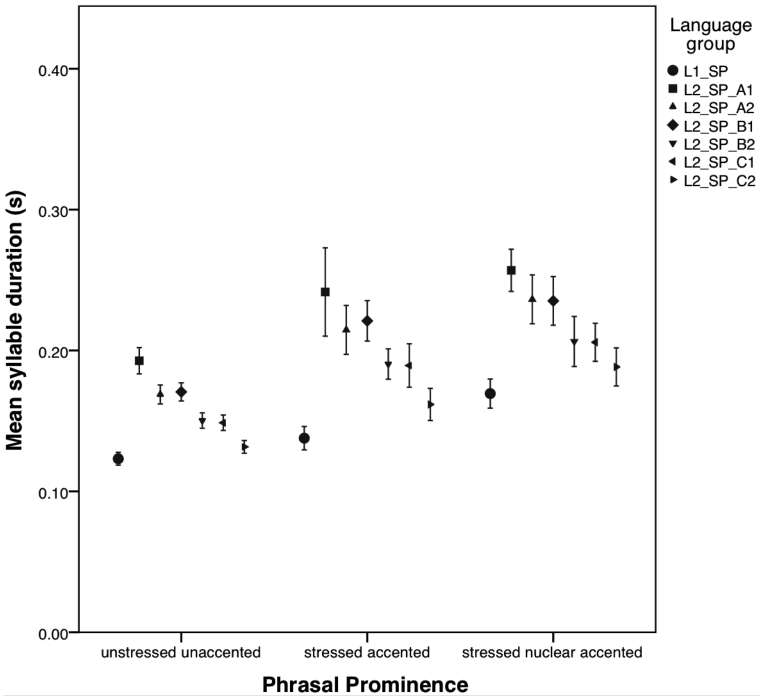


Figure 7. Mean syllable duration (in milliseconds) in L1 Spanish and DLS of all proficiency levels, separated by Phrasal Prominence condition. CV sentences only.

4

Concerning final lengthening, pairwise comparisons show that for non-final syllables the DLS_C2 and DLS_C1 are not significantly different from the L1 Spanish, and for ip-final and IP-final the DLS_C2 are not significantly different from the target L1 Spanish, see **Table 3**. This effect is once again influenced by syllable structure, as making the same comparisons within the CV condition reveals that the three highest proficiency levels are comparable to the L1 target.

Table 3 p-*values of pairwise comparisons between the L1 Spanish (*N *= 5) and all DLS groups (*N *= 30) for Phrasal Position, separated by syllable structure*

Pairwise comparison	All sentences			CV sentences only		
	Non-final	ip-final	IP-final	Non-final	ip-final	IP-final
L1 Spanish – DLS_A1	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p < .001$
L1 Spanish – DLS_A2	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p = .009$	$p = .005$
L1 Spanish – DLS_B1	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p = .014$	$p = .023$
L1 Spanish – DLS_B2	$p = .038$	$p = .001$	$p < .001$	$p = .127$	$p = .773$	$p = .061$
L1 Spanish – DLS_C1	$p = .083$	$p < .001$	$p = .002$	$p = .252$	$p = .396$	$p = .639$
L1 Spanish – DLS_C2	$p = .467$	$p = .305$	$p = 1.000$	$p = .913$	$p = 1.000$	$p = 1.000$

Examination of syllable durations for the different Phrasal Position conditions within each Language Group reveals that the three most proficient DLS groups show a similar pattern as the L1 Spanish in which syllable durations are longer when syllables precede a prosodic boundary. Conversely, the values of the three lowest proficiency groups coincide more with the L1 Dutch, corroborating the presence of transfer effects in L2 rhythm acquisition, see **Figures 8 and 9**, and **Appendix Table A7**.

Finally, the joint effect of Phrasal Position and Phrasal Prominence on syllable durations is examined by inspecting the mean syllable durations for all Phrasal Prominence conditions within the separate Phrasal Position conditions. This reveals that both factors interact systematically within each Language Group: within each Phrasal Position condition accentual lengthening increases as syllables are more prominent in the utterance, while increasing syllable durations are also observed between Phrasal Position conditions, see **Appendix Table A8.**

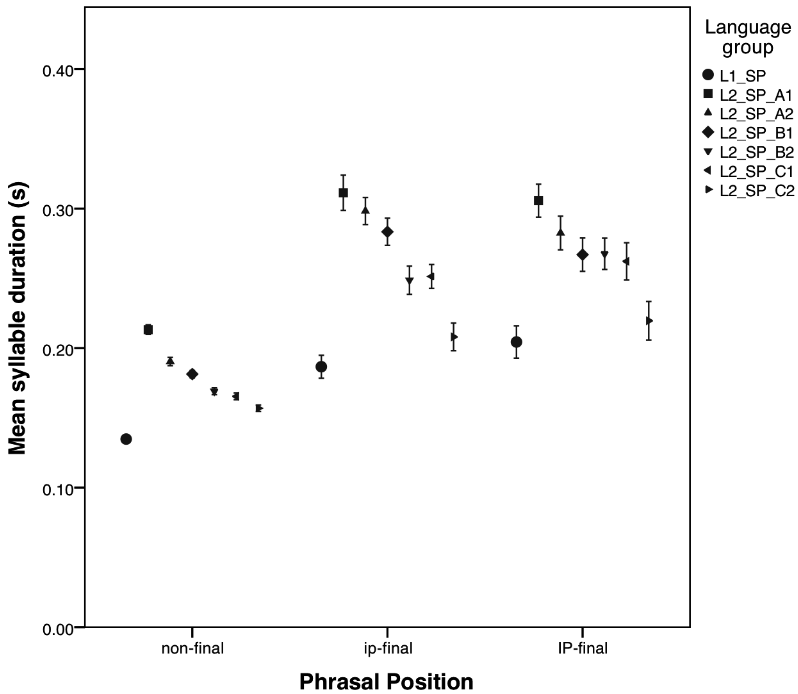


Figure 8. Mean syllable duration (in milliseconds) in L1 Spanish and DLS of all proficiency levels, separated by Phrasal Position condition. All sentences.

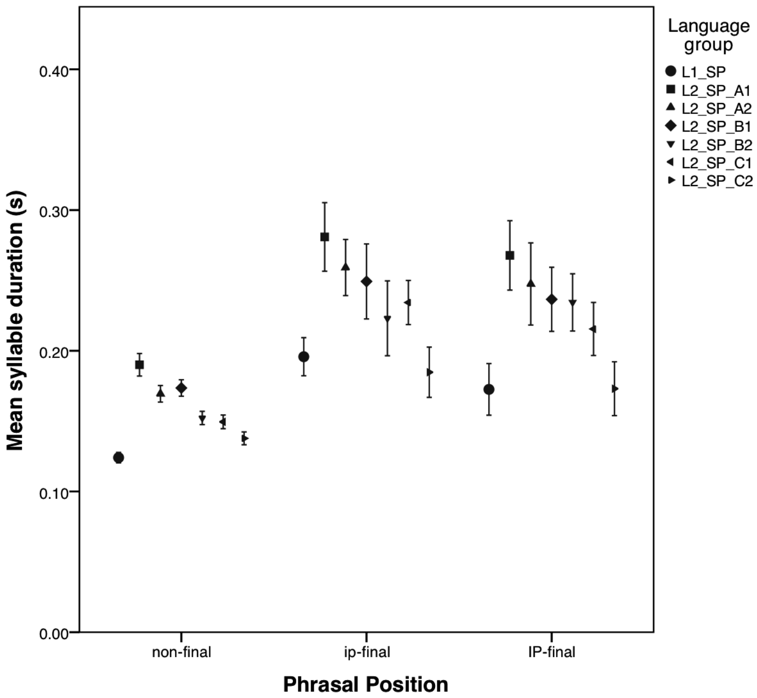


Figure 9. Mean syllable duration (in milliseconds) in L1 Spanish and DLS of all proficiency levels, separated by Phrasal Position condition. CV sentences only.

4.5.3 L1 Dutch versus SLD

To compare the SLD to the L1 speakers of Dutch, a GLMM analysis was performed with syllable duration as the response variable, and Language Group (7 levels: L1 Dutch, SLD_A1, SLD_A2, SLD_B1, SLD_B2, SLD_C1, SLD_C2), Syllable Structure (as above), Phrasal Prominence (as above), and Phrasal Position (as above) as fixed factors. The analysis reveals significant main effects for all fixed factors and significant interactions for all relevant combinations, except for the interaction between Language Group and Phrasal Prominence (see **Appendix Table A9** for all main effects and interactions). Pairwise comparisons between Language Groups overall reveal that although the SLD progressively approach target syllable durations as their proficiency increases, all of the SLD groups still differ significantly from the L1 Dutch (p-values from $p < .001$ to $p = .028$). Crucially, this appears completely due to the syllable structure of the utterances, because when comparing the SLD to the L1 Dutch within the CV condition, the L1 Dutch values do not differ significantly from the SLD values for all proficiency levels (*p*-values from $p = .089$ to $p = 1.000$). Turning to Phrasal Prominence first, the results show that all of the SLD groups differ significantly from the target L1 Dutch for both unstressed and unaccented syllables, and stressed and nuclear accented syllables, see **Table 4**. In the stressed and accented condition, only the SLD_A1 group differs significantly from the L1 Dutch. However, this effect is highly susceptible to the syllable structure of the utterance, as making the same comparisons within the CV condition reveals that all of the SLD groups for all Phrasal Prominence conditions are comparable to the L1 target. Examination of the syllable durations of the different Phrasal Prominence conditions within all Language Groups reveals that both SLD and L1 Dutch show a similar pattern in which syllable durations are longer as syllables are more prominent within an utterance, see **Figures 10 and 11**, and **Appendix Table A10**.

Table 4 p-*values of pairwise comparisons between the L1 Spanish (*N *= 5) and all DLS groups (*N *= 30) for Phrasal Position, separated by syllable structure*

Pairwise comparison	All sentences			CV sentences only		
	Unstressed, unaccented	Stressed, accented	Stressed, nuclear accented	Unstressed, unaccented	Stressed, accented	Stressed, nuclear accented
L1 Dutch – SLD_A1	$p < .001$	$p = .030$	$p < .001$	$p = 1.000$	$p = .583$	$p = 1.000$
L1 Dutch – SLD_A2	$p < .001$	$p = .135$	$p < .001$	$p = .298$	$p = .860$	$p = .476$
L1 Dutch – SLD_B1	$p < .001$	$p = 1.000$	p=.005	$p = 1.000$	$p = 1.000$	$p = 1.000$
L1 Dutch – SLD_B2	$p = .004$	$p = .976$	$p = .006$	$p = 1.000$	$p = 1.000$	$p = 1.000$
L1 Dutch – SLD_C1	$p < .001$	$p = .097$	$p < .001$	$p = 1.000$	$p = .836$	$p = .927$
L1 Dutch – SLD_C2	$p = .006$	$p = .742$	$p = .002$	$p = 1.000$	$p = 1.000$	$p = 1.000$

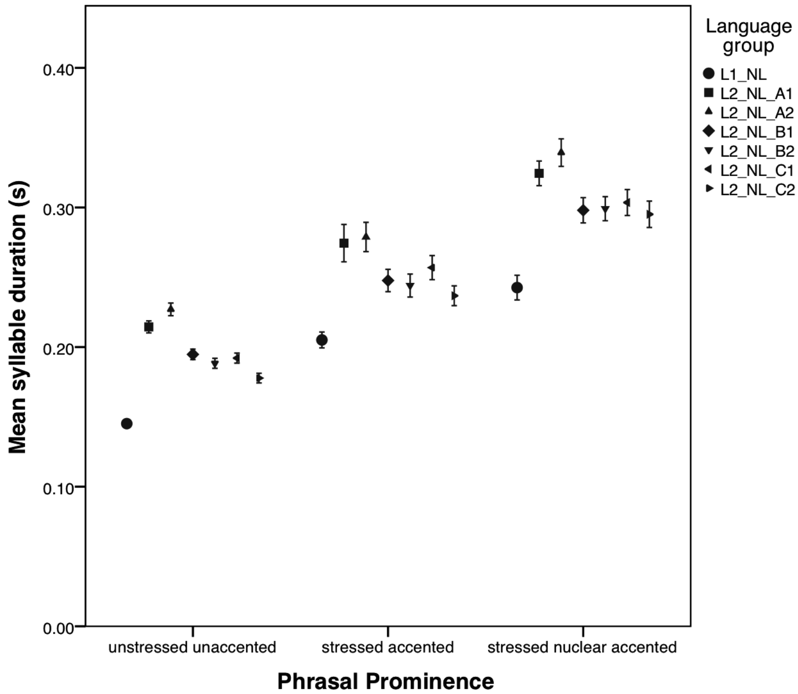


Figure 10. Mean syllable duration (in milliseconds) in L1 Dutch and SLD of all proficiency levels, separated by Phrasal Prominence condition. All sentences.

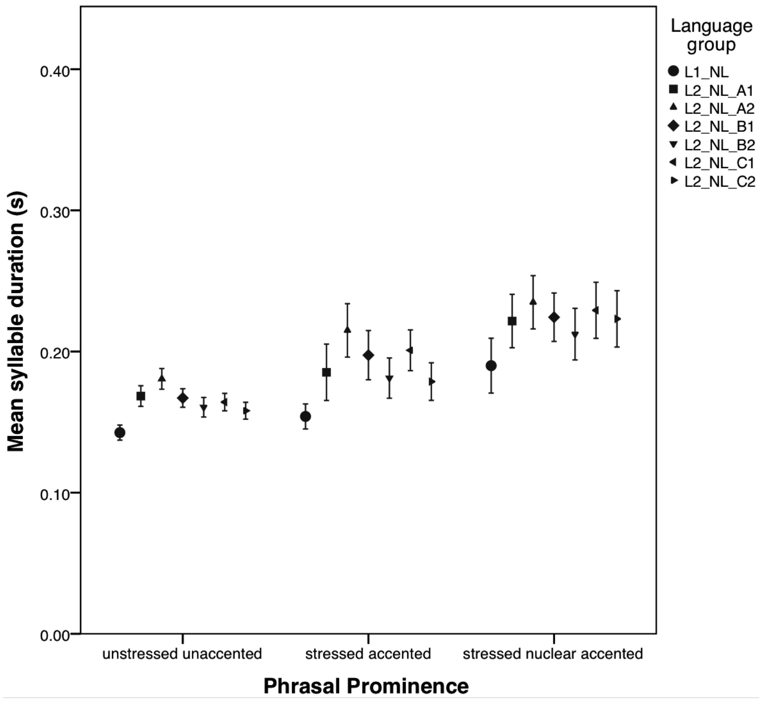


Figure 11. Mean syllable duration (in milliseconds) in L1 Dutch and SLD of all proficiency levels, separated by Phrasal Prominence condition. CV sentences only.

Regarding final lengthening, pairwise comparisons between the different Language Groups within the three Phrasal Position conditions show that for non-final syllables all SLD groups differ significantly from the L1 Dutch, see **Table 5** on the next page. However, for ip-final syllables the SLD_B1, SLD_B2 and SLD_C2 groups are comparable to the L1 Dutch and for the IP-final syllables this is the case for the DLS_B2 and DLS_C1 groups. This effect is again largely due to syllable structure, as identical comparisons in the CV condition reveal that almost all SLD groups are no longer significantly different from the L1 Dutch.

Table 5 *p-values of pairwise comparisons between the L1 Dutch (N = 5) and all SLD groups (N = 30) for Phrasal Position, separated by syllable structure*

Pairwise comparison	All sentences			CV sentences only		
	Non-final	ip-final	IP-final	Non-final	ip-final	IP-final
L1 Dutch – SLD_A1	$p < .001$	$p = .040$	$p < .001$	$p = 1.000$	$p = 1.000$	$p = .066$
L1 Dutch – SLD_A2	$p < .001$	$p = .008$	$p < .001$	$p = .010$	$p = 1.000$	$p = .545$
L1 Dutch – SLD_B1	$p = .003$	$p = .481$	$p = .014$	$p = .487$	$p = 1.000$	$p = 1.000$
L1 Dutch – SLD_B2	$p = .027$	$p = .481$	$p = .058$	$p = 1.000$	$p = 1.000$	$p = 1.000$
L1 Dutch – SLD_C1	$p < .001$	$p < .001$	$p = .306$	$p = .364$	$p = 1.000$	$p = 1.000$
L1 Dutch – SLD_C2	$p = .050$	$p = .339$	$p = .008$	$p = 1.000$	$p = 1.000$	$p = 1.000$

Examination of the syllable durations of the different Phrasal Position conditions within each Language Group reveals that all SLD groups show a similar pattern as the L1 Dutch, in which syllable durations are longer when syllables precede a boundary, either within an utterance or at its end, see **Figures 12 and 13**, and **Appendix Table A11**.

Examining the joint effect of Phrasal Position and Phrasal Prominence on syllable durations by inspection of the mean syllable durations for all Phrasal Prominence conditions within the separate Phrasal Position conditions reveals that both factors interact systematically within each Language Group: within each Phrasal Position condition accentual lengthening increases as syllables are more prominent in the utterance, while increasing syllable durations are also shown between Phrasal Position conditions, see **Appendix Table A12**. Accentual lengthening and final lengthening appear to contribute equally to the differences found between the L1 Dutch and the different SLD groups, especially when controlling for syllable structure. When only analyzing the CV sentences, all SLD groups appear to be fully on target in their syllable duration production, however when diversifying syllable structure (consequently making it more typical of L1 Dutch) it becomes rather more difficult to discern a logical pattern in the SLD productions.

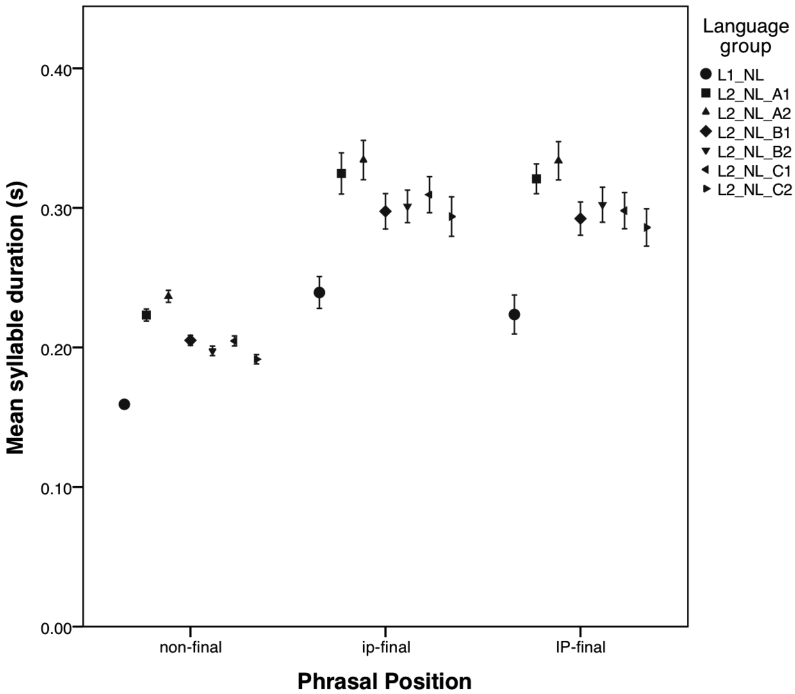


Figure 12. Mean syllable duration (in milliseconds) in L1 Dutch and SLD of all proficiency levels, separated by Phrasal Position condition. All sentences.

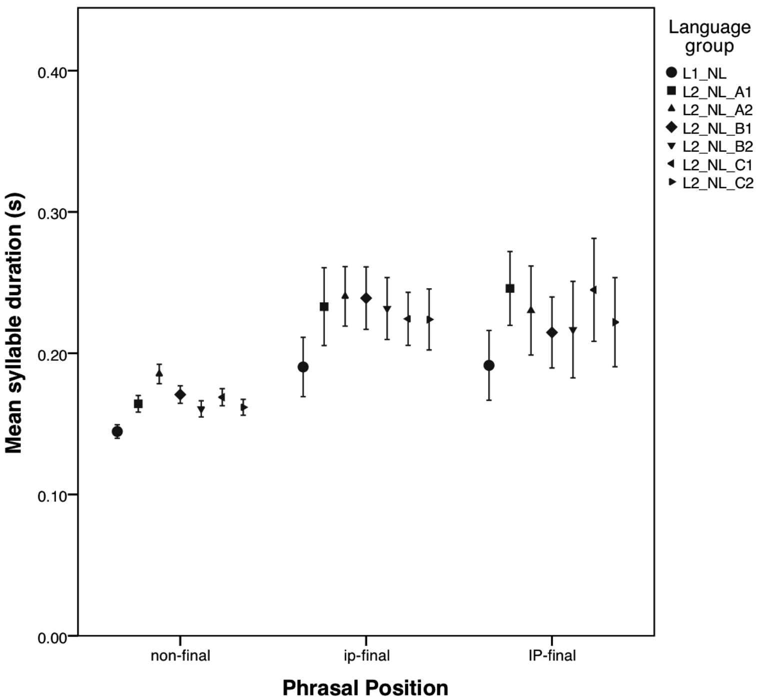


Figure 13. Mean syllable duration (in milliseconds) in L1 Dutch and SLD of all proficiency levels, separated by Phrasal Position condition. CV sentences only.

4.6 Discussion and conclusion

The current study investigated whether L2 learning direction affects the successful attainment of speech rhythm by DLS and SLD. Based on the MDH, we hypothesized that DLS would be more successful at approaching their target than SLD, because both rhythm as a whole, and its correlates syllable structure and lengthening effects, are more marked in Dutch than in Spanish. Overall, our results indeed show that learning direction influences L2 rhythm acquisition: our analyses reveal a different development for DLS than SLD. Comparing the two groups, we can conclude that DLS show a more systematic development towards their target, and more successful attainment of an overall rhythm pattern that coincides with the one produced by L1 Spanish speakers. Thus, our results support our hypothesis and corroborate prior work (Ordin & Polyanskaya, 2015; Rasier, 2006).

However, our results do not allow for a complete disentanglement between speech rhythm and syllable structure complexity: our lengthening analyses revealed different acquisition processes for DLS and SLD. The DLS systematically approach L1 values in all lengthening conditions until attaining target-like values, generally at the highest proficiency level for all sentences, and at an intermediate to advanced level for the CV sentences only. Conversely, SLD of all proficiency levels are completely on target in the CV sentences only, but show no systematic attainment in the analyses including all sentences. Therefore, it seems unlikely that the insignificant difference between the L1 Dutch and the least proficient SLD in the CV sentences is completely due to a perfectly produced speech rhythm by the SLD. Not only do these results show that learning direction influences L2 development, they also suggest that rhythm acquisition by SLD is substantially affected by their difficulties at producing utterances with more complex and/or closed syllable structures: When syllables are more complex and predominantly closed, the SLD are unable to reach target syllable durations, yet when syllables are predominantly open and have a simple CV structure, target patterns appear attainable. In this sense, L2 rhythm acquisition resembles L1 rhythm development in which physical output constraints related to consonant (cluster) production also affect target-like rhythm production (Ordin & Polyanskaya, 2014; Payne et al., 2012).

Similar to Li and Post (2014), our study shows that L2 rhythm acquisition (like L1 rhythm acquisition, see Post & Payne, 2018) is a multisystemic process that requires the simultaneous attainment of several language-specific features, both phonotactic and prosodic. Crucially, depending on the learning direction, some of these features may be more challenging than others. Gradient properties, such as accentual and final lengthening, seem challenging for both DLS and SLD. Yet other, more categorical, characteristics (e.g., syllable structure constraints) appear substantially more difficult to acquire for SLD than DLS. In addition, between-speaker variability may also influence the acquisition process. While we matched our participants to the best of our ability based on their language experience and proficiency and included subject as a random factor in our statistical analyses, individual differences in the L2 acquisition process tend to be substantial (Ellis, 2004) and some factors, such as motivation and language aptitude, could not be taken into account. Especially studies on the multisystemic nature of L2 prosody acquisition would benefit from careful participant selection, as variation across individual speakers might occur in all "systems" and thus be magnified even more. Our study thus reinforces the need for L2 acquisition theories and models that allow for predictions based on the multisystemic nature of L2 prosody acquisition and that accommodate the inclusion of learning direction, as well as other speaker-based characteristics, as a relevant factor.

Moreover, our results are relevant pedagogically, as they demonstrate that adequate segment production is a prerequisite for successful rhythm attainment. The acquisition of suprasegmentals is often overlooked in educational programs, since they are difficult to manipulate consciously, and highly context-dependent. Conversely, the correct pronunciation of segments usually receives considerable attention. On its own, this might not be a bad practice, as the current research suggests that training in this area may also lead to more successful rhythm production. Interestingly, recent work by Polyanskaya, Ordin and Busà (2017) suggests that the relative contribution of segmental characteristics and timing patterns to the assessment of accentedness differ as a factor of the proficiency level of the L2 learner. In other words, while the incorrect pronunciation of segmental properties might contribute more to accentedness in speech produced by less proficient L2 learners, deviance in speech rate and rhythmic patterns could become more salient as L2 learners become more proficient. Future research might therefore be dedicated to production studies investigating this further, as well as to perception studies that may confirm both the effect of deviance in different phonetic areas and in speech by learners of different proficiency levels on judgments by L1 speakers.

Future research could also address the effect of segmental pronunciation training on rhythm acquisition in different developmental stages, in addition to the effect of learning direction for other prosodic features, like lexical stress. Furthermore, since rhythm is related to several language-specific features, the current study could be extended by similar analyses for different L1-L2 combinations. Aside from follow-up studies on L2 production, the effect of (in)correct L2 rhythm production, perhaps in combination with other prosodic features, on L1 perception might be investigated.

Notes

1 Dauer (1983) actually proposes a continuum ranging from less to more stress-timed. In order to maintain the terminology used in previous studies on speech rhythm, we will continue to define the end points of a rhythmic continuum as 'syllable-timed' and 'stress-timed' and use these labels to categorize languages upon the continuum, though we agree that stress-timedness is a gradient, not categorical, feature.

2 For a discussion of the suitability of rhythm metrics for this purpose, see Arvaniti (2009, 2012) and Wiget, White, Schuppler, Grenon, Rauch and Mattys (2010). The latter also present useful recommendations for researchers concerning the use of rhythm metrics in empirical studies.

3 Dutch also makes extensive use of vowel reduction (Koopmans-Van Beinum, 1980), while Spanish does so very little (Delattre, 1969; Hualde, 2005).

4 While comparing the L1 and L2 data of the DLS and SLD might be felicitous in minimizing the effect of individual variability, the L1 data of the L2 learners could not serve as a baseline in the current study as it has been shown that prosodic transfer from the L2 to the L1 occurs in advanced L2 learners (e.g., Mennen, 2004; Van Maastricht et al., 2016a), while it is unknown whether it also occurs in less proficient speakers. To make equal comparisons between all L2 learners and a typical target baseline, it was deemed more suitable to use typical speakers of the L1.

5 This was further investigated by performing two Chi-square analyses (one per language), which show that the number of open syllables differs significantly between the three syllable structure conditions, χ^2 (2, N = 566) = 53.17, p < .001 for Dutch, and χ^2 (2, N = 497) = 96.66, p < .001, for Spanish.

6 While some studies, like the current research, prefer to measure accentual and final lengthening separately, others used one combined lengthening measure (e.g., Li and Post, 2014).

7 To facilitate coding, the last two tiers were coded numerically. In the fifth tier, containing the phrasal position coding, '0' stands for 'non-final', '2' stands for 'ip-final', and '3' corresponds to 'IP-final'

syllables. In the sixth tier, which contains the phrasal prominence coding, '0' stands for 'unaccented and unstressed' syllables, '2' corresponds to 'stressed and accented', and '3' to 'stressed and nuclear accented' syllables.

References

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh, United Kingdom: Edinburgh University Press.

Allen, G. D., & Hawkins, S. (1978). The development of phonological rhythm. In A. Bell, & J. Hooper (Eds.), *Syllables and segments* (pp. 172–185). Amsterdam, The Netherlands: North-Holland Publishing.

Archibald, J. (1994). A formal model of learning L2 prosodic phonology. *Second Language Research*, *10*, 215–240.

Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica*, *66*, 46–63.

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, *40*, 351–373.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278.

Beckman, M., Edwards, J., 1994. Articulatory evidence for differentiating stress categories, phonological structure and phonetic form. In: P.A. Keating (Ed.), *Papers in Laboratory Phonology III* (pp. 7–33). Cambridge, United Kingdom: Cambridge University Press.

Best, C. T. (1995). A direct-realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 171–204). Timonium, Maryland, United States: York Press.

Boersma, P., & Weenink, D. (2010). Praat: Doing phonetics by computer (Version 6.0.04) [Computer software]. Retrieved from http://www.fon.hum.uva.nl/praat/

Bolinger, D. L. (1965). Pitch accent and sentence rhythm. In I. Abe, & T. Kanekiyo (Eds.), *Forms of English: Accent, morpheme, order* (pp. 139–180). Cambridge, Massachusetts, United States: Harvard University Press.

Booij, G. (1995). *The phonology of Dutch*. Oxford, United Kingdom: Clarendon Press.

Bunta, F., & Ingram, D. (2007). The acquisition of speech rhythm by bilingual Spanish- and English-speaking 4- and 5-year-old children. *Journal of Speech, Language, and Hearing Research*, *50*(4), 999–1014.

Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica, 57*, 3–16.

Cambier-Langeveld, T., Nespor, M., & van Heuven, V. J. (1997). The domain of final lengthening in production and perception in Dutch. In *Proceeding of the European Conference on Speech Communication and Technology, 2*, 931-934.

Cambier-Langeveld, T., & Turk, A. E. (1999). A cross-linguistic study of accentual lengthening: Dutch vs. English. *Journal of Phonetics*, *27*(3), 255-280.

4

Carter, P. M. (2005). Quantifying rhythmic differences between Spanish, English, and Hispanic English. In R.S. Gess, & E. Rudin (Eds.), *Theoretical and experimental approaches to Romance linguistics: Selected papers from the 34th Linguistic Symposium on Romance Languages, Salt Lake City, March 2004* (pp. 63-75). Amsterdam, The Netherlands: John Benjamins Publishing Company.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge, United Kingdom: Cambridge University Press.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalysed. *Journal of Phonetics*, *11*, 51–62.

Dauer, R. M. (1987). Phonetic and phonological components of language rhythm. In *Proceedings of the 11th International Congress of Phonetic Sciences* (pp. 447-450). Tallinn, Estonia: Academy of Sciences.

Delattre, P. (1966). A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics in Language Teaching*, *4*, 183–198.

Delattre, P. (1969). An acoustic and articulatory study of vowel reduction in four languages. *International Review of Applied Linguistics for Language Teaching, 7*, 295-325.

Dellwo, V. (2006). Rhythm and speech rate: A variation coefficient for ΔC. In P. Karnowski, & I. Szigeti (Eds.), *Language and language processing: Proceedings of the 38th Linguistic Colloquium, Piliscsaba 2003* (pp. 231–241). Frankfurt am Main, Germany: Peter Lang.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning 27*(2), 315–330.

Eckman, F. (2008). Typological markedness and second language phonology. In J. G. Hansen Edwards, & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 95–115). Philadelphia, Pennsylvania, United States: John Benjamins.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*(4), 551-585.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, Maryland, United States: York Press.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology*, *7*, 515–546.

Grabe, E., Watson, I., & Post, B. (1999). The acquisition of rhythmic patterns in English and French. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 1201–1204). San Francisco, California, United States: University of California.

Gut, U. (2003). Non-native speech rhythm in German. In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2437-2440). Barcelona, Spain: Universitat Autonomá de Barcelona.

Gut, U. (2009). *Non-native speech: A corpus-based analysis of phonological and phonetic properties of L2 English and German.* Frankfurt am Main, Germany: Peter Lang.

Hartsuiker, R. J. (2002). The addition bias in Dutch and Spanish phonological speech errors: The role of structural context. *Language and Cognitive Processes*, *17*(1), 61-96.

Hualde, J. I. (2005). *The sounds of Spanish*. Cambridge, United Kingdom: Cambridge University Press.

IBM Corporation. (2015). SPSS Statistics (version 22) [Computer software]. Available from http://www.spss.com/software/statistics/

Koopmans-Van Beinum, F. J. (1980). *Vowel contrast reduction: An acoustic and perceptual study of Dutch vowels in various speech conditions*. Amsterdam, The Netherlands: Academische Pers.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Levelt, C., & Van de Vijver, R. (2004). Syllable types in cross-linguistic and developmental grammars. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Fixing priorities: Constraints in phonological acquisition* (pp. 204-218). Cambridge, United Kingdom: Cambridge University Press.

Li, A., & Post, B. (2014). L2 acquisition of prosodic properties of speech rhythm. *Studies in Second Language Acquisition*, *36*(2), 223-255.

Lubbers, M., & Torreira, F. (2015). Praatalign: An interactive Praat plug-in for performing phonetic forced alignment (version 1.9b) [Computer software]. Retrieved from https://github.com/dopefishh/Praatalign

Mairano, P., & Romano, A. (2010). Un confronto tra diverse metriche ritmiche usando Correlatore ('A comparison of different rhythm metrics using Correlatore', in Italian). In S. Schmid, M. Schwarzenbach, & D. Studer (Eds.), *La dimensione temporale del parlato*, *Proceedings of the 5th National AISV Congress* (pp. 79-100). Torriana, Italy: EDK.

Navarro Tomás, T. (1966). *Estudios de fonologia española* ('Studies on Spanish phonology', in Spanish). New York, New York, United States: Las Américas.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human perception and performance*, *24*(3), 756-766.

Ordin, M., & Polyanskaya, L. (2014). Development of timing patterns in first and second languages. *System*, *42*, 244-257.

Ordin, M., & Polyanskaya, L. (2015). Acquisition of speech rhythm in a second language by learners with rhythmically different native languages. *The Journal of the Acoustical Society of America*, *138*(2), 533-544.

Özçelik, Ö. (2016). The Prosodic Acquisition Path Hypothesis: Towards explaining variability in L2 acquisition of phonology. *Glossa: A Journal of General Linguistics*, *1*(1), 1-48.

Payne, E., Post, B., Prieto, P., Vanrell, M., & Astruc, L. (2012). Measuring child rhythm. *Language and Speech*, *55*, 203–229.

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America*, *32*, 693–703.

Pike, K. L. (1945). *The intonation of American English*. Ann Arbor, Michigan, United States: University of Michigan Press.

Polyanskaya, L., & Ordin, M. (2015). Acquisition of speech rhythm in first language. *Journal of the Acoustical Society of America, 138*(3), 199-204.

Polyanskaya, L., Ordin, M., & Busà, M. (2017). Relative salience of speech rhythm and speech rate on perceived foreign accent in a second language. *Language and Speech, 60*(3), 333-355.

Post, B., & Payne, E. (2018). Speech rhythm in development: What is the child acquiring? In: P. Prieto, & N. Esteve-Gibert (Eds.), *The development of prosody in first language acquisition*. Amsterdam, The Netherlands: John Benjamins.

Prieto, P., Vanrell, M., Astruc, L., Payne, E., & Post, B. (2012). Phonotactic and phrasal properties of speech rhythm: Evidence from Catalan, English, and Spanish. *Speech Communication*, *54*, 681–702.

Ramus, F., Dupoux, E., & Mehler, J. (2003). The psychological reality of rhythm classes: Perceptual studies. In M. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 337–342). Barcelona, Spain: Universitat Autonomá de Barcelona.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *The Journal of the Acoustical Society of America*, *105*(1), 512-521.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*, 265–292.

Rasier, L. (2006). *Prosodie en vreemdetaalverwerving: accentdistributie in het Frans en in het Nederlands als vreemde taal* ('Prosody and second language acquisition: Accent distribution in French and Dutch as a foreign language', in Dutch) (Doctoral dissertation). Université catholique de Louvain, Louvain, Belgium.

Schiller, N. O., Meyer, A. S., Baayen, R. H., & Levelt, W. J. (1996). A comparison of lexeme and speech syllables in Dutch. *Journal of Quantitative Linguistics*, *3*(1), 8-28.

Schmidt, E., & Post, B. (2015). The development of prosodic features and their contribution to rhythm production in simultaneous bilinguals. *Language and Speech*, *58*(1), 24-47.

Valls Ferrer, M. (2011). *The development of oral fluency and rhythm during a study abroad period* (Doctoral dissertation). Universitat Pompeu Fabra, Barcelona, Spain.

Van Zon, M. (1997). *Speech processing in Dutch: A cross-linguistic approach* (Doctoral dissertation). Katholieke Universiteit Brabant, Tilburg, The Netherlands.

White, L., & Mattys, S. L. (2007a). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, *35*, 501–522.

White, L., & Mattys, S. L. (2007b). Rhythmic typology and variation in first and second languages. In P. Prieto, J. Mascaró, & M.-J. Solé (Eds.), *Segmental and prosodic issues in romance phonology* (Current issues in linguistic theory series) (pp. 237–257). Amsterdam, The Netherlands: John Benjamins.

Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., & Mattys, S. (2010). How stable are acoustic metrics of contrastive speech rhythm? *Journal of the Acoustical Society of America, 127*, 1559–1569.

Appendix

Table A1 *Speaker characteristics*

L1	L2	Gender	Age	Proficiency	*N*
Dutch	-	Female	29 and 59 years	Native	2
Dutch	-	Male	25, 31 and 38 years	Native	3
Spanish	-	Female	34 and 50 years	Native	2
Spanish	-	Male	32, 38 and 48 years	Native	3
Dutch	Spanish	Female	40, 48 and 57 years	A1	3
Dutch	Spanish	Male	57 and 59 years	A1	2
Dutch	Spanish	Female	20, 21 and 21 years	A2	3
Dutch	Spanish	Male	20 and 21 years	A2	2
Dutch	Spanish	Female	19, 22 and 22 years	B1	3
Dutch	Spanish	Male	24 and 33 years	B1	2
Dutch	Spanish	Female	20, 22 and 23 years	B2	3
Dutch	Spanish	Male	22 and 38 years	B2	2
Dutch	Spanish	Female	32, 36, 36 and 61 years	C1	4
Dutch	Spanish	Male	45 years	C1	1
Dutch	Spanish	Female	36, 44 and 45 years	C2	3
Dutch	Spanish	Male	29 and 57 years	C2	2
Spanish	Dutch	Female	22 years	A1	1
Spanish	Dutch	Male	36, 37, 37 and 38 years	A1	4

Table A1 *Continued*

L1	L2	Gender	Age	Proficiency	*N*
Spanish	Dutch	Female	19, 23 and 49 years	A2	3
Spanish	Dutch	Male	28 and 29 years	A2	2
Spanish	Dutch	Female	36 and 42 years	B1	2
Spanish	Dutch	Male	30, 47 and 54 years	B1	3
Spanish	Dutch	Female	27, 27 and 34 years	B2	3
Spanish	Dutch	Male	37 and 38 years	B2	2
Spanish	Dutch	Female	34, 43, 50 and 66 years	C1	4
Spanish	Dutch	Male	36 years	C1	1
Spanish	Dutch	Female	44, 46, 47and 50 years	C2	4
Spanish	Dutch	Male	40 years	C2	1

Sentences used in the reading task. The number in parentheses represents the total number of syllables, followed by the number of orthographic words.

Dutch

CV sentences

(1) De mama van Nadine komt uit Malaga. (12, 7)
(2) De Bacardi uit Havana is van goede origine. (16, 8)
(3) De mama van Susana is een gezellige lerares. (16, 8)
(4) De baby heeft de hele papaja en kiwi opgegeten. (17, 9)
(5) Panama en Cuba liggen in Midden-Amerika. (15, 7)

CVC sentences

(6) Roos had haarbanden gekocht in Londen met haar vriendin Annabel. (17,10)
(7) De woonboot van Wim is niet heel ruim, maar wel zeer comfortabel. (16, 12)
(8) Marcel had z'n hoed verloren met het rennen naar de veerboot. (16, 11)
(9) De wedstrijd van de voetbalclub was niet in het sportcomplex. (15, 10)
(10) Dokter van der Wal had wel zesendertig patiënten behandeld die dag. (19, 11)

Mixed sentences

(11) De boze demonstranten raakten slaags met de politie. (15, 8)
(12) Het omstreden artikel zorgde voor heel wat opschudding. (15, 8)
(13) De prinses had kramp in haar hand van het lintjes doorknippen. (15, 11)
(14) De uitgever spande een proces aan tegen de schrijver. (15, 9)
(15) De dader werd helaas bij gebrek aan bewijs vrijgesproken. (16, 9)
(16) Het belang van milieubewustheid wordt steeds vaker ingezien. (16, 8)
(17) Zij heeft voor alles altijd een psychologische verklaring. (16, 8)
(18) Onze laatste aanwinst is een prachtige antieke sofa. (16, 8)
(19) Op het ijs stond een kraampje met chocolademelk en stroopwafels. (17, 10)
(20) Delegaties uit meer dan twintig landen komen naar dit congres. (17, 10)
(21) Het lawaai van de machines maakte elk gesprek onmogelijk. (17, 9)
(22) Een gevoel van enorme opluchting maakte zich van hem meester. (17, 10)
(23) In tegenstelling tot zijn broer heeft hij altijd van schaken gehouden. (18, 11)
(24) Beeldend kunstenaars doen vaak hun inspiratie op in grote steden. (18, 10)
(25) Onder grote spanning vertonen de meeste mensen hun ware aard. (18, 10)
(26) Door het uitvallen van de microfoons was de toespraak onverstaanbaar. (18, 10)
(27) In die dierentuin is voor de eerste maal een pandabeertje geboren. (19, 11)
(28) De favoriete werkplek van de schrijfster was een oude villa aan zee. (19, 12)
(29) Niemand heeft ooit kunnen achterhalen waar het geld terechtgekomen is. (19, 10)
(30) Het was die dag zo heet dat de toeristen spontaan in de fontein sprongen. (19, 14)

Spanish

CV sentences

(1) La madre de Susana es de Badalona. (13, 7)
(2) La banana de Guatemala es de buena calidad. (16, 8)
(3) La madre de Susana es una buena profesora. (16, 8)
(4) El logo de la fábrica se diseñó en Cataluña. (16, 9)
(5) Canadá y Perú no están en Centroamérica. (13, 7)

CVC sentences

(6) Los donuts de Ámsterdam son realmente internacionales. (17, 7)
(7) Las naranjas de Londres no son las más dulces del mundo. (16,11)
(8) Los mangos del Brasil y Ceilán son de calidad extra. (16, 10)
(9) El meeting del club de tenis no fue en el parking del club. (15, 13)
(10) El doctor Frankenstein es un monstruo sentimental e internacional. (19, 9)

Mixed sentences

(11) Dio su último concierto en el teatro municipal. (15, 8)
(12) Para eso necesitaremos mucho más dinero. (15, 6)
(13) Los padres se acercaron del niño sin hacer ruido. (15, 9)
(14) El niño se levantó temprano para ver el sol. (15, 9)
(15) Los niños salen todos los días a la misma hora. (16, 10)
(16) Reportan inundaciones graves en la primavera. (16, 6)
(17) La radio anunció esta noticia el miércoles pasado. (16, 8)
(18) El ladrón se fue con los pendientes de oro de mi madre. (16, 12)
(19) Encontrar un empleo no es fácil en el contexto actual. (17, 10)
(20) Los artistas siempre fueron atraídos por las ciudades. (17, 8)
(21) Mucha gente vino a celebrar la victoria con nosotros. (17, 9)
(22) Los bancos cierran particularmente temprano el viernes. (17, 7)
(23) Los vecinos de mis abuelos son gente súper agradable. (18, 9)
(24) Mis padres leyeron la noticia en el diario *El Periódico*. (18, 10)
(25) Llueve durante todo el año en los países tropicales. (18, 9)
(26) La corriente ecológica creció bastante en la clase media. (18, 9)
(27) El presupuesto del ministerio de la cultura bajó mucho. (19, 9)
(28) No entendí nada del libro que tú me prestaste hace dos semanas. (19, 12)
(29) La reconstrucción de la ciudad empezó la semana pasada. (19, 9)
(30) El director declaró que la situación estaba bajo control. (19, 9)

Table A2 *Lengthening: overview of relevant main effects and interactions for L1 speakers (N = 10)*

Fixed factor(s)	Effect
Language Group	**$F(1, 4950) = 21.85, p < .001$**
Syllable Structure	**$F(2, 4950) = 16.31, p < .001$**
Phrasal Prominence	**$F(2, 4950) = 156.49, p < .001$**
Phrasal Position	**$F(2, 4950) = 109.12, p < .001$**
Language Group*Syllable Structure	**$F(2, 4950) = 24.45, p < .001$**
Language Group*Phrasal Prominence	**$F(2, 4950) = 27.96, p < .001$**
Language Group*Phrasal Position	$F(2, 4950) = 1.04, p = .353$
Language Group*Syllable Structure*Phrasal Prominence	**$F(8, 4950) = 10.10, p < .001$**
Language Group*Syllable Structure*Phrasal Position	**$F(8, 4950) = 7.44, p < .001$**
Language Group*Phrasal Prominence*Phrasal Position	**$F(5, 4950) = 6.72, p < .001$**

Note. Bold characters are used to mark significant effects.

Table A3 *Mean syllable durations (standard deviations) in seconds produced by L1 speakers of Dutch (N = 5) and Spanish (N = 5), separated per Phrasal Prominence condition, for all sentences and CV sentences only*

Language Group	Phrasal Prominence	*M* syllable duration (*SD*)	
		All sentences	*CV sentences only*
L1 Dutch	unstressed, unaccented	.15 (.05)	.14 (.04)
	stressed, accented	.21 (.06)	.15 (.04)
	stressed, nuclear accented	.24 (.08)	.19 (.07)
L1 Spanish	unstressed, unaccented	.13 (.05)	.12 (.04)
	stressed, accented	.15 (.05)	.14 (.03)
	stressed, nuclear accented	.18 (.06)	.17 (.03)

Table A4 *Mean syllable durations (standard deviations) in seconds produced by L1 speakers of Dutch (*N *= 5) and Spanish (*N *= 5), separated per Phrasal Position condition, for all sentences and CV sentences only*

Language Group	Phrasal Position	*M* syllable duration (*SD*)	
		All sentences	*CV sentences only*
L1 Dutch	non-final	.16 (.06)	.14 (.04)
	ip-final	.24 (.08)	.19 (.04)
	IP-final	.22 (.09)	.19 (.07)
L1 Spanish	non-final	.13 (.05)	.12 (.03)
	ip-final	.19 (.05)	.20 (.03)
	IP-final	.20 (.07)	.17 (.04)

Table A5 *Lengthening: overview of relevant main effects and interactions for DLS (*N *= 30) in comparison to L1 speakers of Spanish (*N *= 5)*

Fixed factor(s)	Effect
Language Group	**$F(6, 17759) = 15.74, p < .001$**
Syllable Structure	**$F(2, 17759) = 17.23, p = .001$**
Phrasal Prominence	**$F(2, 17759) = 368.79, p < .001$**
Phrasal Position	**$F(2, 17759) = 472.91, p < .001$**
Language Group*Syllable Structure	**$F(12, 17759) = 2.71, p = .001$**
Language Group*Phrasal Prominence	**$F(12, 17759) = 5.38, p < .001$**
Language Group*Phrasal Position	**$F(12, 17759) = 8.54, p < .001$**
Language Group*Syllable Structure*Phrasal Prominence	$F(28, 17759) = .60, p = .952$
Language Group*Syllable Structure*Phrasal Position	**$F(28, 17759) = 4.79, p < .001$**
Language Group*Phrasal Prominence*Phrasal Position	**$F(16, 17759) = 2.24, p = .003$**

Note. Bold characters are used to mark significant effects.

Table A6 *Mean syllable durations (standard deviations) in seconds produced by L1 speakers of Spanish (N = 5) and DLS of all proficiency levels (N = 30), separated per Phrasal Prominence condition, all sentences*

Language Group	Phrasal Prominence	*M* syllable duration (*SD*)
L1 Spanish	unstressed, unaccented	.13 (.05)
	stressed, accented	.15 (.05)
	stressed, nuclear accented	.18 (.06)
DLS_A1	unstressed, unaccented	.22 (.09)
	stressed, accented	.25 (.08)
	stressed, nuclear accented	.28 (.10)
DLS_A2	unstressed, unaccented	.19 (.08)
	stressed, accented	.23 (.08)
	stressed, nuclear accented	.25 (.09)
DLS_B1	unstressed, unaccented	.18 (.07)
	stressed, accented	.22 (.07)
	stressed, nuclear accented	.25 (.08)
DLS_B2	unstressed, unaccented	.17 (.06)
	stressed, accented	.20 (.06)
	stressed, nuclear accented	.23 (.08)
DLS_C1	unstressed, unaccented	.17 (.06)
	stressed, accented	.20 (.06)
	stressed, nuclear accented	.22 (.07)
DLS_C2	unstressed, unaccented	.15 (.06)
	stressed, accented	.18 (.06)
	stressed, nuclear accented	.21 (.07)

Table A7 *Mean syllable durations (standard deviations) in seconds produced by L1 speakers of Spanish (*N = *5) and DLS of all proficiency levels (*N = *30), separated per Phrasal Position condition, all sentences*

Language Group	Phrasal Position	*M* syllable duration (*SD*)
L1 Spanish	non-final	.13 (.05)
	ip-final	.19 (.05)
	IP-final	.20 (.07)
DLS_A1	non-final	.21 (.08)
	ip-final	.31 (.10)
	IP-final	.31 (.10)
DLS_A2	non-final	.19 (.07)
	ip-final	.30 (.08)
	IP-final	.28 (.09)
DLS_B1	non-final	.18 (.06)
	ip-final	.28 (.08)
	IP-final	.27 (.09)
DLS_B2	non-final	.17 (.06)
	ip-final	.25 (.07)
	IP-final	.27 (.08)
DLS_C1	non-final	.17 (.06)
	ip-final	.25 (.07)
	IP-final	.26 (.09)
DLS_C2	non-final	.16 (.05)
	ip-final	.21 (.07)
	IP-final	.22 (.09)

Table A8 *Mean syllable durations (standard deviations) in seconds produced by L1 speakers of Spanish (N = 5) and DLS of all proficiency levels (N = 30), separated per Phrasal Position condition, all sentences*

Language Group	Phrasal Position	Phrasal Prominence	*M* syllable duration (*SD*)
L1 Spanish	non-final	unstressed, unaccented	.13 (.04)
		stressed, accented	.15 (.05)
		stressed, nuclear accented	.16 (.05)
	ip-final	unstressed, unaccented	.18 (.05)
		stressed, nuclear accented	.21 (.04)
	IP-final	unstressed, unaccented	.19 (.07)
		stressed, nuclear accented	.24 (.05)
DLS_A1	non-final	unstressed, unaccented	.20 (.07)
		stressed, accented	.25 (.08)
		stressed, nuclear accented	.25 (.08)
	ip-final	unstressed, unaccented	.29 (.09)
		stressed, nuclear accented	.37 (.09)
	IP-final	unstressed, unaccented	.29 (.10)
		stressed, nuclear accented	.35 (.10)
DLS_A2	non-final	unstressed, unaccented	.18 (.06)
		stressed, accented	.23 (.08)
		stressed, nuclear accented	.22 (.07)
	ip-final	unstressed, unaccented	.28 (.08)
		stressed, nuclear accented	.34 (.07)
	IP-final	unstressed, unaccented	.27 (.09)
		stressed, nuclear accented	.33 (.07)

Table A8 *Continued*

Language Group	Phrasal Position	Phrasal Prominence	*M* syllable duration (*SD*)
DLS_B1	non-final	unstressed, unaccented	.17 (.05)
		stressed, accented	.22 (.07)
		stressed, nuclear accented	.22 (.06)
	ip-final	unstressed, unaccented	.26 (.08)
		stressed, nuclear accented	.34 (.07)
	IP-final	unstressed, unaccented	.25 (.08)
		stressed, nuclear accented	.32 (.08)
DLS_B2	non-final	unstressed, unaccented	.15 (.05)
		stressed, accented	.20 (.06)
		stressed, nuclear accented	.20 (.06)
	ip-final	unstressed, unaccented	.23 (.07)
		stressed, nuclear accented	.30 (.06)
	IP-final	unstressed, unaccented	.26 (.08)
		stressed, nuclear accented	.31 (.07)
DLS_C1	non-final	unstressed, unaccented	.15 (.05)
		stressed, accented	.20 (.06)
		stressed, nuclear accented	.20 (.06)
	ip-final	unstressed, unaccented	.23 (.06)
		stressed, nuclear accented	.30 (.06)
	IP-final	unstressed, unaccented	.25 (.09)
		stressed, nuclear accented	.29 (.06)

Table A8 *Continued*

Language Group	Phrasal Position	Phrasal Prominence	*M* syllable duration (*SD*)
DLS_C2	non-final	unstressed, unaccented	.15 (.05)
		stressed, accented	.18 (.06)
		stressed, nuclear accented	.19 (.06)
	ip-final	unstressed, unaccented	.19 (.06)
		stressed, nuclear accented	.25 (.07)
	IP-final	unstressed, unaccented	.21 (.02)
		stressed, nuclear accented	.25 (.07)

Table A9 *Lengthening: overview of relevant main effects and interactions for SLD (N = 30) in comparison to L1 speakers of Dutch (N = 5)*

Fixed factor(s)	Effect
Language Group	**$F(6, 17208) = 6.81, p < .001$**
Syllable Structure	**$F(2, 17208) = 75.80, p < .001$**
Phrasal Prominence	**$F(2, 17208) = 834.82, p < .001$**
Phrasal Position	**$F(2, 17208) = 173.25, p < .001$**
Language Group*Syllable Structure	**$F(12, 17208) = 3.79, p < .001$**
Language Group*Phrasal Prominence	$F(12, 17208) = 1.12, p = .340$
Language Group*Phrasal Position	**$F(12, 17208) = 1.91, p = .029$**
Language Group*Syllable Structure*Phrasal Prominence	**$F(28, 17208) = 7.78, p < .001$**
Language Group*Syllable Structure*Phrasal Position	**$F(28, 17208) = 4.10, p < .001$**
Language Group*Phrasal Prominence*Phrasal Position	**$F(21, 17208) = 6.79, p < .001$**

Note. Bold characters are used to mark significant effects.

Table A10 *Mean syllable durations (standard deviations) in seconds produced by L1 speakers of Dutch (*N *= 5) and SLD of all proficiency levels (*N *= 30), separated per Phrasal Prominence condition, all sentences*

Language Group	Phrasal Prominence	*M* syllable duration (*SD*)
L1 Dutch	unstressed, unaccented	.15 (.05)
	stressed, accented	.21 (.06)
	stressed, nuclear accented	.24 (.08)
SLD_A1	unstressed, unaccented	.21 (.09)
	stressed, accented	.27 (.10)
	stressed, nuclear accented	.32 (.11)
SLD_A2	unstressed, unaccented	.23 (.09)
	stressed, accented	.28 (.09)
	stressed, nuclear accented	.34 (.12)
SLD_B1	unstressed, unaccented	.19 (.08)
	stressed, accented	.25 (.08)
	stressed, nuclear accented	.30 (.11)
SLD_B2	unstressed, unaccented	.19 (.07)
	stressed, accented	.24 (.08)
	stressed, nuclear accented	.30 (.10)
SLD_C1	unstressed, unaccented	.19 (.07)
	stressed, accented	.26 (.08)
	stressed, nuclear accented	.30 (.11)
SLD_C2	unstressed, unaccented	.18 (.07)
	stressed, accented	.24 (.07)
	stressed, nuclear accented	.30 (.10)

Table A11 *Mean syllable durations (standard deviations) in seconds produced by L1 speakers of Dutch (*N *= 5) and SLD of all proficiency levels (*N *= 30), separated per Phrasal Position condition, all sentences*

Language Group	Phrasal Position	*M* syllable duration (*SD*)
L1 Dutch	non-final	.16 (.06)
	ip-final	.24 (.08)
	IP-final	.22 (.09)
SLD_A1	non-final	.22 (.09)
	ip-final	.32 (.11)
	IP-final	.32 (.11)
SLD_A2	non-final	.24 (.10)
	ip-final	.33 (.11)
	IP-final	.33 (.13)
SLD_B1	non-final	.21 (.08)
	ip-final	.30 (.09)
	IP-final	.29 (.11)
SLD_B2	non-final	.20 (.08)
	ip-final	.30 (.09)
	IP-final	.30 (.11)
SLD_C1	non-final	.20 (.08)
	ip-final	.31 (.11)
	IP-final	.30 (.10)
SLD_C2	non-final	.19 (.08)
	ip-final	.29 (.10)
	IP-final	.29 (.10)

Table A12 *Mean syllable durations (standard deviations) in seconds produced by L1 Dutch (N = 5) and SLD of all proficiency levels (N = 30), separated per Phrasal Position and Phrasal Prominence combination, all sentences*

Language Group	Phrasal Position	Phrasal Prominence	*M* syllable duration (*SD*)
L1 Dutch	non-final	unstressed, unaccented	.14 (.05)
		stressed, accented	.20 (.06)
		stressed, nuclear accented	.21 (.07)
	ip-final	unstressed, unaccented	.21 (.07)
		stressed, nuclear accented	.29 (.07)
	IP-final	unstressed, unaccented	.19 (.06)
		stressed, nuclear accented	.31 (.09)
SLD_A1	non-final	unstressed, unaccented	.20 (.08)
		stressed, accented	.27 (.10)
		stressed, nuclear accented	.27 (.10)
	ip-final	unstressed, unaccented	.26 (.08)
		stressed, nuclear accented	.40 (.09)
	IP-final	unstressed, unaccented	.27 (.09)
		stressed, nuclear accented	.39 (.10)
SLD_A2	non-final	unstressed, unaccented	.21 (.08)
		stressed, accented	.28 (.09)
		stressed, nuclear accented	.29 (.11)
	ip-final	unstressed, unaccented	.27 (.08)
		stressed, nuclear accented	.40 (.10)
	IP-final	unstressed, unaccented	.29 (.12)
		stressed, nuclear accented	.41 (.12)

Table A12 *Continued*

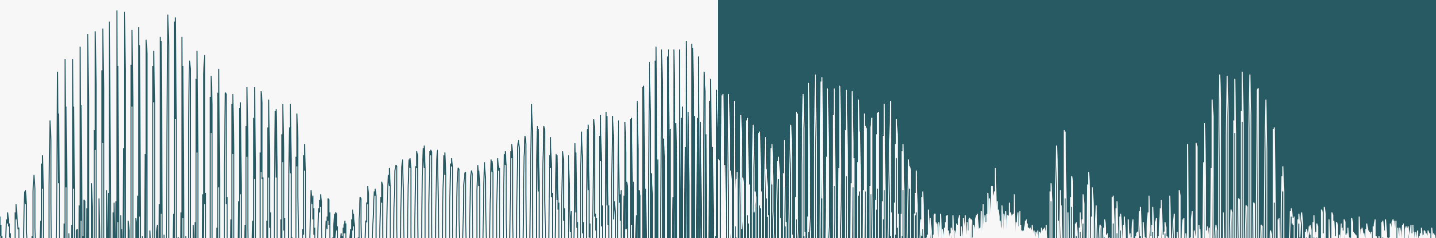
Language Group	Phrasal Position	Phrasal Prominence	*M* syllable duration (*SD*)
SLD_B1	non-final	unstressed, unaccented	.18 (.07)
		stressed, accented	.25 (.08)
		stressed, nuclear accented	.26 (.10)
	ip-final	unstressed, unaccented	.25 (.08)
		stressed, nuclear accented	.35 (.08)
	IP-final	unstressed, unaccented	.25 (.09)
		stressed, nuclear accented	.36 (.10)
SLD_B2	non-final	unstressed, unaccented	.17 (.06)
		stressed, accented	.24 (.08)
		stressed, nuclear accented	.25 (.09)
	ip-final	unstressed, unaccented	.25 (.08)
		stressed, nuclear accented	.36 (.07)
	IP-final	unstressed, unaccented	.25 (.09)
		stressed, nuclear accented	.37 (.09)
SLD_C1	non-final	unstressed, unaccented	.18 (.06)
		stressed, accented	.25 (.08)
		stressed, nuclear accented	.26 (.09)
	ip-final	unstressed, unaccented	.26 (.09)
		stressed, nuclear accented	.38 (.10)
	IP-final	unstressed, unaccented	.25 (.08)
		stressed, nuclear accented	.37 (.09)

Table A12 *Continued*

Language Group	Phrasal Position	Phrasal Prominence	*M* syllable duration (*SD*)
SLD_C2	non-final	unstressed, unaccented	.16 (.06)
		stressed, accented	.24 (.07)
		stressed, nuclear accented	.26 (.09)
	ip-final	unstressed, unaccented	.25 (.08)
		stressed, nuclear accented	.36 (.09)
	IP-final	unstressed, unaccented	.24 (.09)
		stressed, nuclear accented	.37 (.08)

5

rhythm perception



5

The interplay of prosodic cues in the L2

How the speech rate, intonation, and rhythm of speech by Spanish learners of Dutch contributes to L1 Dutch perceptions of accentedness and comprehensibility*

Abstract

This study investigates the relative contribution of L2 intonation, rhythm, and speech rate to L1 perceptions of accentedness and comprehensibility. To this end, the intonation, rhythm, and speech rate of a L1 speaker of Dutch was transferred onto the segmental string of four Spanish learners of Dutch, resulting in eight conditions that reflect all possible combinations of these prosodic cues. Our results show that improving the prosody of L2 learners positively generally affects L1 perceptions, but not to the same extent for both perception measures: Concerning accentedness, Dutch listeners were influenced equally by intonation and speech rate manipulations, but not by the rhythm manipulation. No interaction effects were found between prosodic cues. Conversely, comprehensibility ratings were affected most by intonation, and to a lesser degree by speech rate, and several interaction effects were found between all the prosodic features. Our study reaffirms the importance of differentiating between different aspects of perception and provides insight into those features that are most likely to affect each L1 perception type.

* **This chapter is based on:** Van Maastricht, L., Zee, T., Krahmer, E., & Swerts, M. (2018). The interplay of prosodic cues in a second language: How the speech rate, intonation, and rhythm of speech by Spanish learners of Dutch contributes to L1 Dutch perceptions of accentedness and comprehensibility. Under revision at *Language and Speech*.

5.1 Introduction

Adult foreign language (L2) learners are generally unable to attain the same native-like accent as mother tongue (L1) speakers. Previous studies have shown that adult L2 speech typically shows characteristics of the speakers' L1, both on a segmental and suprasegmental level. Segmentally, adult learners tend to have difficulty distinguishing between L2 phonemes if they do not represent a phonological contrast in their L1, and their production of these L2 phonemes is often also off-target. For example, Dutch learners of Spanish usually find it difficult to discriminate between the Spanish [r] (as in *carro*, wagon) and [ɾ] (as in *caro*, expensive), because while Dutch is characterized by a large variation in rhotics, none of these contrast phonemically (Booij, 1995). L2 learners also have difficulty acquiring sounds that supersede the phonemic level, such as the intonation patterns and rhythm of the L2. For the learner group in the present study, that is, Spanish learners of Dutch, it has been shown that transfer effects occur from the L1 to the L2 in both the production of phrasal prominence (Chapter 2 of this dissertation) and the production of rhythm (Chapter 4).

While many previous studies investigated these transfer effects in L2 production (e.g., Li & Post, 2014; Mennen, 2004), the effects of L1 transfer in L2 speech on perception by L1 speakers remains relatively under-investigated. Especially lacking are perception studies from a multisystemic perspective in which the effect of transfer in multiple cues is taken into account simultaneously, while also distinguishing between different types of L1 perception of L2 speech, for instance related to perceived foreign accent, perceived difficulty of comprehension, actual processing of the speech. In other words, while it is both theoretically and intuitively clear that there is a perceivable difference between speech produced by L1 and L2 speakers and that the use of prosodic cues from the L1 in the L2 contributes to this difference (e.g., Munro & Derwing, 2001; Chapter 3 of this dissertation), very few studies have actually investigated the relative contribution of these cues. To the best of our knowledge, previous work on the perception of L2 speech has usually been limited to one or two prosodic properties at a time. For instance, prior studies have shown that the production of target-like speech rate, intonation, and rhythm affects how L1 listeners perceive L2 speakers (e.g., Ito & Speer, 2008; Munro & Derwing, 2001; Quené & Van Delft, 2010), but they have not investigated how combinations of these parameters influence those perceptions. Combining several prosodic features in one design yields stimuli that are more similar to actual L2 speech and enables analyses of the interaction between these features in perception.

Moreover, there are different types of L1 perceptions of L2 speech, depending on whether the focus is on the degree of perceived foreign accent (accentedness), the reported ease of understanding of L2 speech (comprehensibility), or measures of actual processing of L2 speech by L1 listeners, such as reaction times (intelligibility). Most prior work on the contribution of prosody to L1 perceptions of L2 speech only investigated one aspect of perception, usually accentedness. However, previous studies have shown that atypical prosodic features affect accentedness, comprehensibility, and intelligibility in different ways and to varying degrees. For example, in Chapter 3 of this dissertation, we reported mean accentedness ratings on a 9-point scale ranging from 2.4 to 8.7 depending on the proficiency of the speaker, while mean comprehensibility ratings only varied from 5.9 to 8.5. This difference in perceptual measures is not only found when a specific prosodic cue with a specific communicative function is manipulated (e.g., pitch accent distributions used to mark focus as in Chapter 3 of this dissertation), but also when broader prosodic characteristics of speech, such as intonation, are measured as a whole, without distinguishing between their different functions and realizations (e.g., Munro & Derwing, 1999). As both the perception (e.g., Anderson-Hsieh, Johnson & Koehler, 1992; Caspers & Horłoza, 2012) and the production (e.g., Li & Post, 2014; Chapter 4 of this dissertation) of L2 speech are known to result from the interplay between various L1 and L2 features, it is relevant to determine what the relative and cumulative effects are of multiple prosodic properties on the aforementioned types of L1 perception of L2 speech, which is the aim of the current study.

To this end, speech by Spanish learners of Dutch has been manipulated using speech resynthesis techniques in order to create the stimuli for a rating task performed by L1 speakers of Dutch. Before turning to the specifics of our experiment, we briefly discuss the typological differences between Spanish and Dutch concerning the prosodic features that are relevant to the current study and we review the previous studies that examined the separate effects of atypical intonation, rhythm and speech rate in the L2, using accentedness and/or comprehensibility as dependent variables.

5.2 Theoretical background

5.2.1 Prosodic differences: Spanish versus Dutch

Turning to rhythmic differences between Dutch and Spanish first, Spanish was traditionally characterized as a syllable-timed language in which the duration of all syllables was assumed to be similar, and Dutch was taken to be a stress-timed language in which the duration of the intervals between stressed syllables was considered to be similar (Abercrombie, 1967; Pike, 1945). Since then, it was shown in acoustic studies that there is no evidence for such isochrony and that this categorical distinction cannot be maintained (e.g., Bolinger, 1965; Dauer, 1983). Moreover, recent studies have shown that Spanish and Dutch differ in certain phonetic, phonotactic, and prosodic properties that in turn influence our perception of their speech rhythm. Phonetic features include for instance vowel reduction, which is employed more extensively in Dutch than in Spanish (Delattre, 1969; Hualde, 2005; Koopmans-Van Beinum, 1980). The same holds for the prosodic properties of final and accentual lengthening, that is, the marking of utterance boundaries and prominence by means of lengthening syllable duration, which are also employed significantly more in Dutch (Cambier-Langeveld, Nespor & Van Heuven, 1997; Cambier-Langeveld & Turk, 1999) than in Spanish (Delattre, 1966; Prieto et al., 2012; Chapter 4 of this dissertation). Finally, Dutch and Spanish differ in their syllable complexity: Spanish is characterized by a preference for simple, open syllables consisting predominantly of a consonant and a vowel, while Dutch predominantly uses more complex, closed syllables (Hartsuiker, 2002; Navarro Tomás, 1966). The combination of these language-specific properties has been shown to correlate with perceptual speech rhythm differences between both languages and the idea of a rhythm continuum of "stress-timedness" (Dauer, 1983), on which Dutch and Spanish lie on opposite ends (Nazzi, Bertoncini & Mehler, 1998; Prieto et al., 2012). They have also been corroborated by analyses using rhythm metrics (e.g., White & Mattys, 2007). In Chapter 4 of this dissertation, we showed that

Spanish learners of Dutch have difficulty producing a speech rhythm that is typical of L1 Dutch, even at an advanced level, unless the utterance predominantly consists of the CV syllables that are typical of their L1. In other words, their inability to successfully produce consonant clusters at the segmental level contributes to their non-native rhythm production.

Concerning intonation, Dutch and Spanish are similar on several accounts, for example, they both do not use pitch contours contrastively at the lexical level (i.e., they are not tone languages), but they do employ lexical stress contrastively: in Dutch *voorkomen* means 'to occur', while *voorkomen* means 'to prevent', and *hablo* means 'I speak', while *habló* means 'he spoke' in Spanish. In addition, both languages use phrasal intonation to express pragmatic meanings such as surprise and sarcasm or to mark the illocutionary force of an utterance (Beckman, Díaz-Campos, Tevis McGory & Morgan, 2002; Gussenhoven, 2005). One important aspect in which Dutch and Spanish are known to differ though, is the use of pitch accents to mark focus (Chapter 2 of this dissertation). While in Dutch pitch accents tend to be placed on new or contrastive information and given information is usually deaccented, irrespective of the position of this element in the utterance, in Spanish pitch accents are typically placed on all prosodic words, and nuclear stress is applied to intonational phrase-final elements, which generally entails that word order changes are used to ensure that this intonational phrase-final element contains the new or contrastive information. Chapter 2 of this dissertation showed that both Spanish learners of Dutch and Dutch learners of Spanish are unable to produce completely target-like pitch accent distributions in their L2 as a result of transfer effects from the L1, even at an advanced level. This is congruent with other studies showing that L2 learners generally have difficulties acquiring native-like intonation (e.g., Mennen, 2004; Ramírez Verdugo, 2003).

With respect to speech rate, to our knowledge, no direct comparisons have been made between Dutch and Spanish. A review of the relevant literature suggests that in spontaneous speech, Dutch and Spanish do not differ from each other in this respect: Dutch speech rate has been reported to be in the order of 4.23 syllables per second in Verhoeven, De Pauw, and Kloots (2004), while Spanish speech rate has been reported as 4.24 syllables per second in De Johnson, O'Connell, and Sabin (1979). However, more important for the current study is the fact that the speech rate of L1 and L2 speech tends to differ, irrespective of the L1 or L2 of the speaker. It has been shown repeatedly that L2 speech is generally slower than L1 speech (see Guion, Flege, Liu & Yeni-Komshian, 2000, and the references therein), which means that manipulating this prosodic aspect of speech might influence L1 listeners' perceptions of its accentedness. In summary, while Dutch and Spanish differ prosodically, it is not clear how these different suprasegmental cues interact in perceptions by L1 speakers. In the subsequent section, the existing literature on the individual effects of rhythm, intonation and prosody on different aspects of perception is reviewed.

5.2.2 L1 perceptions of L2 speech: accentedness, comprehensibility, and intelligibility

This section reviews the different aspects of L1 perception of L2 speech that have been studied in previous work and it defines our concepts of L1 perception of accentedness, comprehensibility and intelligibility, respectively. Derwing and Munro defined accentedness as the extent to which "an L2 accent differs from the variety of speech commonly spoken in the community" (Derwing & Munro, 2005: 385). In this sense, it can be interpreted as the opposite of nativeness, that is, "the degree to which a speaker sounds like a native speaker of a particular language" (Edmunds, 2009: 2). Derwing and Munro defined comprehensibility as "the listener's perception of the degree of difficulty encountered when trying to understand an utterance" (Derwing & Munro, 2005: 385). Lastly, intelligibility has been defined as "the extent to which the speaker's intended utterance is actually understood by a listener" (Derwing & Munro, 2005: 385). It is especially important to distinguish clearly between the latter two measures, as both concepts are known to overlap in previous work. Comprehensibility is an offline measure, which reflects listeners' *evaluations* of the effort required to process a speech fragment, usually measured on a semantic differential (e.g., Anderson-Hsieh, Johnson & Koehler, 1992, who incidentally use the term intelligibility to refer to what we would call comprehensibility). Conversely, intelligibility is an online measure that reflects actual processing and is generally operationalized by means of transcription tasks (e.g., Derwing & Munro, 1997) or accuracy scores in combination with reaction times (e.g., Chapter 3 of this dissertation) or eye tracking (e.g., Ito & Speer, 2008).

Unfortunately, many studies on the effects of atypical prosody in L2 speech are limited to only one or at most two of these aspects of L1 perceptions, which means that a more complete overview of those aspects of L1 perception of L2 speech that are affected by a specific language feature is often lacking. There are some notable exceptions: for example, Saito, Trofimovich and Isaacs (2016) performed a multiple regression analysis to examine the contribution of segmental errors, word stress errors, intonation, and speech rate (as well as other, non-phonetic factors) to comprehensibility and accentedness ratings. Their research on utterances produced by Japanese learners of English showed that all areas of pronunciation correlated significantly with both measures, with no differences in strength of the association for comprehensibility, while accentedness was more strongly associated with segmental errors than with intonation and speech rate, suggesting that these two dependent variables indeed measure different aspects of perception. Furthermore, in Chapter 3 of this dissertation, we measured L1 Dutch perceptions of accentedness, comprehensibility, and intelligibility with respect to the intonation of utterances produced by Spanish learners of Dutch. We found that while accentedness and comprehensibility were both affected by incorrect pitch accent placement to mark focus, albeit to a different extent (comprehensibility was affected less strongly than accentedness), intelligibility was not affected at all. This incongruence between different aspects of L1 perceptions of L2 speech was also reported by Winters and O'Brien (2013), who used speech resynthesis to create different conditions in which (1) native intonation and/or syllable durations[1] were transferred onto non-native segments and (2) non-native intonation and/or syllable durations were transferred onto native segments in both English and German. For the stimuli in which L1 prosody was combined with L2 segments, they reported a reduction in perceived accentedness, but also decreased intelligibility, which suggests that the perceptual processes of intelligibility and accentedness in the L2 are not identical.

While some of the studies cited above compare the effect of multiple prosodic cues on aspects of L1 perception of L2 speech, studies examining these prosodic cues in isolation are much more common. For instance, Munro and Derwing (2001) examined the effect of speech rate changes on foreign accent and comprehensibility ratings. They reported that L1 English listeners generally considered typical L2 English with a relatively low speech rate as more accented and less easy to comprehend than L2 English that was somewhat faster than typical L2 speech. However, both very fast and very slow speech were rated as more accented and more difficult to understand than speech with a moderately low or moderately fast speech rate. In order to determine the effect of intonation on foreign accent perception, Van Els and De Bot (1987) flattened the pitch contour in speech by L2 speakers of Dutch with a French, English or Turkish background, as well as L1 speakers of Dutch, while retaining its segmental properties. This manipulation affected the degree of success with which L1 Dutch listeners were able to determine whether they heard an L1 or L2 speaker of Dutch, suggesting that intonational features also contribute to the perception of a foreign accent. Magen's (1998) study using stimuli produced by L2 English speakers with a Hispanic background further corroborates this view as it showed that the use of atypical lexical and phrasal stress affected accentedness ratings as well.

The studies that investigated the contribution of durational (i.e., rhythmic) properties[2] to intelligibility (e.g., Quené & Van Delft, 2010; Tajima, Port, & Dalby, 1997), usually looked at the combined effect of rhythmic and intonational properties on perceived foreign accent: Boula de Mareüil and Vieru Dimulescu (2006) showed that the transferring the melodic and durational characteristics of one language onto the segmental string of a related language (in this case Spanish and Italian) resulted in speech that was more often classified as the language from which the prosody was transferred than as the language from which the segments were taken. Yet combining intonation and rhythm in one holistic manipulation of prosody does not enable a disentanglement of their respective effects. While Ramus and Mehler (1999) did manipulate rhythm and intonation separately, their dependent measure did not truly reflect foreign accentedness: their participants were asked to distinguish between two fictional languages 'Sahatu' and 'Moltec', while actually distinguishing between English and Japanese. They used speech resynthesis techniques to create four conditions in which they preserved either (1) the intonation, rhythm, and segments, (2) the rhythm and intonation, (3) only the intonation, or (4) only the rhythm of English and Japanese utterances. Their results revealed that rhythm was a "necessary and sufficient cue" for French adults when discriminating between these English and Japanese utterances (Ramus & Mehler, 1999: 1).

Presently, only one study combined all three prosodic cues in one design (though using a less specific intonation manipulation than the one we used in the current study) to determine how L2 rhythm, intonation, and speech rate contribute to foreign accent perceptions by L1 speakers: Polyanskaya, Ordin and Busa (2016) manipulated utterances by French learners of English in such a way that (1) the original L2 rhythm was preserved, while controlling for speech rate, (2) the L2 speech rate was maintained, while rhythm was controlled for, and (3) both rhythm and speech rate were preserved. All stimuli were created with the same intonational contour as well as with monotonized pitch. Their results showed that both speech rate and rhythm affected accentedness, but that rhythm did so more strongly than speech rate. Additionally, intonation was shown to boost the perception of minimal rhythmic differences, but to reduce the salience of small differences in speech rate.

5.3 The current study

Thus, while prior research investigated the individual effects of intonation, rhythm, and speech rate on either accentedness or comprehensibility, no study ever combined all three prosodic features in one design, while also measuring their contribution to different aspects of L1 perception of L2 speech. As accentedness ratings tend to range from one extremity of the scale to the other, whereas comprehensibility ratings only occupy part of the scale, we argue that it is important to include both perception measures in one study, using identical stimuli and manipulations for both tasks. As intelligibility requires a different operationalization than accentedness and comprehensibility, which can both be measured on a scale, we chose to focus on the latter two measures of perception in this study. In order to allow for individual manipulations of all three prosodic cues, resynthesis techniques were used to create the stimuli used in the present study. Based on the relevant literature, the following predictions can be made concerning the effect of individual prosodic cues on accentedness and or comprehensibility:

(1) It is predicted that transferring the *speech rate* of a Dutch L1 speaker onto L2 speech by Spanish learners of Dutch will positively affect both accentedness and comprehensibility ratings by L1 Dutch listeners (Munro & Derwing, 2001; Polyanskaya et al., 2016; Saito, Trofimovich & Isaacs, 2016).

(2) Based on Magen (1998), Saito et al. (2016), Van Els and De Bot (1987), and Chapter 3 of this dissertation, it can be predicted that transferring the *intonation* of a Dutch L1 speaker onto L2 speech by Spanish learners of Dutch will positively affect both accentedness and comprehensibility ratings by L1 Dutch listeners. However, Ramus and Mehler (1999) and Polyanskaya et al. (2016) found a negative effect of intonation on both measures. As the studies that found a negative effect either used a different dependent measure (Ramus & Mehler, 1999) or a different way to manipulate intonation (Polyanskaya et al., 2016), we expect that intonation does improve perceptions of accentedness and comprehensibility.

(3) Transferring the *speech rhythm* of a Dutch L1 speaker onto L2 speech by Spanish learners of Dutch will positively affect accentedness ratings by L1 Dutch listeners (Polyanskaya et al., 2016; Ramus & Mehler, 1999). No studies were found that reported an effect of rhythm on comprehensibility.

Due to the varying methods and measures used in previous studies and the lack of relevant data, we will refrain from making specific predictions concerning the interactions between rhythm, intonation and speech rate.

5.4 Method

5.4.1 Participants

60 adults (*M* age = 20.72 years, *SD* = 2.40) performed our rating task: 43 women and 27 men. All were monolingually raised Dutch students or PhD

candidates at Tilburg University, the former participating for course credit and the latter voluntarily. Controlling for L2 proficiency is difficult in the Netherlands, as English, German and French are generally obligatory at the higher levels of the Dutch educational system. However, it was ensured that none of the participants spoke any other L2 than those three, and none of the participants reported speaking French at a higher level than standard high school proficiency. All participants had completed higher secondary education and had done so relatively recently (the oldest participant being 27 years old).

5.4.2 Materials

The experiment was created and presented to the participants in the online survey tool Qualtrics (2017). All utterances that were used in the current study were previously elicited and recorded during a study on L2 rhythm production (Chapter 4 of this dissertation) in which L1 speakers of Dutch and Spanish, as well as adult learners of both languages, read aloud thirty sentences. Ten of these utterances were used in the current perception experiment; they were produced by a L1 speaker of Dutch whose intonation, rhythm, and speech rate functioned as donor material and by four Spanish L2 speakers of Dutch whose segmental strings were used as receiver structures. Of these four Spanish learners of Dutch, two had successfully completed the B1 level course at the Escuela Oficial de Idiomas in Madrid (implying that they were developing from intermediate to upper intermediate users), and the other two had successfully completed the C1 level course at the same institute (suggesting that they were advanced speakers, Council of Europe, 2001). The four Spanish learners of Dutch were all raised monolingually in the Madrid area and, aside from English, none of them spoke any other Germanic language. As small differences naturally occur between speakers, the four speakers chosen for this study were selected based on the syllabic structures that they used in their utterances. These had to be identical to those used by the donor L1 speaker in order to successfully transfer the prosodic features of the L1 speaker onto the segmental string of the L2 speakers. To guarantee comparability across stimuli and to facilitate prosodic manipulations, only material by female speakers was selected for the current perception task.

The selection of the 10 sentences from a larger sample available for all speakers was made by taking into account the syllable structure of the utterances, controlling for sufficient fluency in the production of the sentences by the Spanish learners of Dutch, and by discarding recordings with poorer sound quality. All sentences were typical of the Dutch language regarding both word frequency and syllable structure, that is, all contained a mix of open and closed syllables of varying complexity (for more details, see Chapter 4). A list of the 10 Dutch utterances used in the current perception study including their translation to English is presented in (4)-(13).

(4) *De baby heeft de hele papaya en kiwi opgegeten* (The baby ate the entire papaya and kiwi)

(5) *Roos had haarbanden gekocht in Londen met haar vriendin Annabel* (Roos bought headbands in London with her friend Annabel)

(6) *De woonboot van Wim is niet heel ruim, maar wel zeer comfortabel* (Wim's houseboat is not very large, but it is very comfortable)

(7) *Dokter van der Wal had wel zesendertig patiënten behandeld die dag* (Dr. van der Wal treated thirty-six patients that day)

(8) *De dader werd helaas bij gebrek aan bewijs vrijgesproken* (Unfortunately, the culprit was acquitted due to a lack of evidence)

(9) *Zij heeft voor alles altijd een psychologische verklaring* (She always has a psychological explanation for everything)

(10) *Onze laatste aanwinst is een prachtige, antieke sofa* (Our last purchase is a beautiful, antique sofa)

(11) *Delegaties uit meer dan twintig landen komen naar dit congress* (Delegations from more than twenty countries are coming to this congress)

(12) *Een gevoel van enorme opluchting maakte zich van hem meester* (An enormous feeling of relief came over him)

(13) *Onder grote spanning vertonen de meeste mensen hun ware aard* (Under pressure, most people show their true nature)

The following instructions and 7-point scales for accentedness (14) and comprehensibility (15) (based on Chapter 3 of this dissertation) were presented to the participants (English translations of the original Dutch sentences).

(14) Indicate to which extent the speaker you heard has a foreign accent

No foreign accent – Very strong foreign accent

(15) Indicate to which degree the speaker you heard is easy/difficult to understand

Incomprehensible – Very easy to understand

5.4.3 Prosodic manipulations

In the current experiment, three prosodic parameters were manipulated: speech rate, rhythm, and intonation. These features were extracted from a donor utterance and transferred onto a receiver utterance of the same sentence. A L2 speaker always produced the receiver utterance, while the donor utterances were produced by a L1 speaker for the seven experimental conditions, and by a L2 speaker of the same proficiency level as the receiver for the control condition, see **Table 1**.[3] The control condition was also submitted to the transplantation and resynthesis process, albeit using donor material of the other L2 learner of the same proficiency level, to ensure comparability across conditions. The utterances were prosodically manipulated by modifying the pitch and duration of the speech signal using the PSOLA algorithm (Moulines & Charpentier, 1990) implemented in Praat (Boersma & Weenink, 2016, version 6.0.21). As previous papers have provided in-depth explanations of this process and its use in L2 perception research (e.g., Boula de Mareüil & Vieru-Dimulescu, 2006; Jilka, 2000; Pettorino & Vitale, 2012; Yoon, 2007), this section focuses on its specific application to the present experiment, starting with the preparation of the speech material.

Table 1 *Mean (SD), range, minimum and maximum speech rate in syllables per second for each speaker group, separated by condition*

Condition	Donor Speaker	Recipient Speaker	Transferred features
C1	L1 Spanish	L1 Spanish	Intonation, rhythm, speech rate
C2	L1 Dutch	L1 Spanish	Intonation
C3	L1 Dutch	L1 Spanish	Rhythm
C4	L1 Dutch	L1 Spanish	Speech rate
C5	L1 Dutch	L1 Spanish	Intonation, rhythm
C6	L1 Dutch	L1 Spanish	Intonation, speech rate
C7	L1 Dutch	L1 Spanish	Rhythm, speech rate
C8	L1 Dutch	L1 Spanish	Intonation, rhythm, speech rate

As previously mentioned, successful manipulation of prosody requires that the donor and receiver versions of an utterance contain the same number of syllables and pauses. Hence, only utterances that matched in the number of syllables that were used for each word were selected, as determined by manual word and syllable level annotations. Subsequently, if either version of the remaining sentences did not contain a silent pause that was present in the other version, a very small corresponding pause was inserted in which the amplitude was set to close to zero to ensure that both pauses and speech were transferred accurately. Another requirement is an accurate F0 analysis of both the receiver and the donor utterance, which is needed in order to prevent unnatural sounding shifts in pitch. To that end, optimal values for the time step, maximum frequency, and minimum frequency of each analysis were manually selected. Visual inspection of the resulting pitch marks compared to the periodicity in the waveform and aural inspection of the synthesized pitch served as selection criteria. The result of these steps served as the input to the subsequent manipulation process, as represented in the upper part of **Figure 1**. Depending on the experimental condition, the manipulation step consisted of resynthesizing the receiver utterance using either the speech rate, rhythm, intonation, or any combination of those parameters from the donor utterance, see **Table 1** for the resulting conditions.

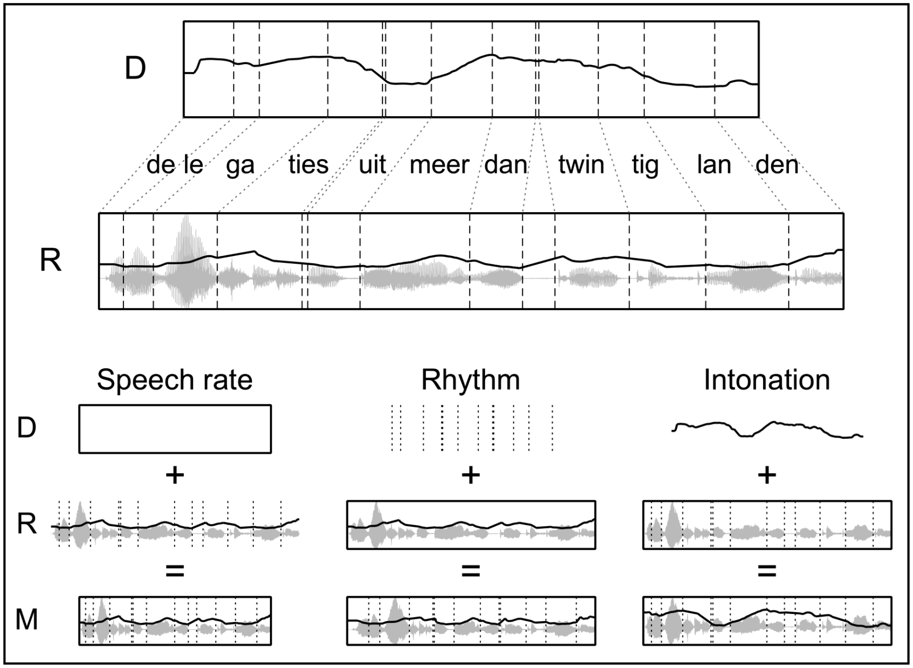


Figure 1. Schematic overview of the manipulation process for part of a stimulus utterance. The length of the rectangular outlines represents total duration of the utterance, the vertical dashed lines represent rhythm, and intonation is represented by the F0 contour. D, R, and M refer to donor, receiver, and manipulated utterance.

The transfer of speech rate was implemented as follows: The total duration of the receiver speech signal was either compressed or stretched to match the duration of the donor signal. The leftmost column of the lower part of **Figure 1** shows how the duration of a donor L1 utterance was used to compress the receiving L2 utterance. Crucially, no changes were made to the relative durations of the syllables, all syllables were compressed using the same ratio. **Table 2** displays the mean (standard deviation), range, minimum and maximum speech rate, measured as syllables per second, per speaker group separated by condition.

Table 2 *Mean (SD), range, minimum and maximum speech rate in syllables per second for each speaker group, separated by condition*

Condition	Speaker Group	Mean (SD)	Range	Minimum	Maximum
C1, C2, C3, C5	L2_intermediate	3.83 (.63)	2.54	3.00	5.54
	L2_advanced	4.08 (.76)	3.00	2.69	5.69
C4, C6, C7, C8	L2_intermediate	5.25 (.66)	2.40	3.87	6.27
	L2_advanced	5.25 (.66)	2.40	3.87	6.27
Donor	L1	5.25 (.68)	2.40	3.87	6.27

As reported in prior work (e.g., Munro & Derwing, 2001 and the references therein), L2 speech was typically articulated more slowly than L1 speech, and an increase in the proficiency level of the L2 speakers coincides with a higher speech rate. In those conditions where the speech rate of the L1 donor is transferred onto the L2 material (i.e., C4, C6, C7, and C8), the speech rate naturally is identical to the L1 rate.

To transfer the L1 rhythm onto the L2 segmental string, each receiver syllable was compressed or stretched so that its duration relative to the total duration of a receiver utterance was identical to the proportional duration of the corresponding donor syllable. By adapting the proportional duration of all syllables in this way, the resulting utterance reflected the final and accentual lengthening patterns typical of L1 Dutch. This is exemplified by the middle column in the bottom part of **Figure 1**. **Table 3** displays the mean (standard deviation), range, minimum and maximum combined accentual and final lengthening ratio (LR), per speaker group, separated by condition.

Table 3 *Mean (SD), range, minimum and maximum combined lengthening ratio per speaker group, separated by condition*

Condition	Speaker Group	Mean (SD)	Range	Minimum	Maximum
C1, C2, C4, C6	L2_intermediate	1.55 (.27)	.86	1.12	1.98
	L2_advanced	1.59 (.28)	.95	1.09	2.04
C3, C5, C7, C8	L2_intermediate	1.64 (.30)	.83	1.20	2.03
	L2_advanced	1.64 (.30)	.83	1.20	2.03
Donor	L1	1.64 (.31)	.83	1.20	2.03

This final measure has been calculated using the formula in (16), in which S_1 is the set of all syllables that are lengthened due to either accentual or final lengthening, and S_2 is its complement, that is, all non-final or non-accented syllables. The upper part of the formula calculates the mean duration of lengthened syllables by taking the sum of the duration (*d*) of all elements (*x*) in S_1, divided by the total number of elements in S_1. This numerator is divided by the denominator, which consists of the same procedure for S_2 and represents the mean duration of syllables that were not lengthened.

(16) $$LR = \frac{\sum_{x_1 \in S_1} d_{x_1} / \overline{\overline{S_1}}}{\sum_{x_2 \in S_2} d_{x_2} / \overline{\overline{S_2}}}$$

Prior work on final and accentual lengthening in Dutch (e.g., Cambier-Langeveld, 1999) and Spanish (e.g., Prieto et al., 2012) showed that L1 Dutch is typically characterized by more extensive final and accentual lengthening than L1 Spanish, and L2 learners generally have trouble attaining target-like production where these prosodic features are concerned (e.g., Li & Post, 2014). The values reported in **Table 3** corroborate this, as they show that lengthening as a factor of prominence or finality marking occurs most extensively in the donor material, and least extensively in the speech of Spanish learners of Dutch at an intermediate level. Proficient Spanish learners of Dutch lengthen more than the intermediate learners do, but do not yet perform on target. In those conditions where the rhythm of the L1 donor is transferred onto the L2 material (i.e., C3, C5, C7, and C8), the combined lengthening ratio of these manipulated stimuli naturally is identical to the L1 ratio.

Intonation was also transferred on a syllable-by-syllable basis. The pitch contour of each donor syllable was made to fit the corresponding receiver syllable, as shown in the right-most column in the lower part of **Figure 1**. To retain the receiving speaker's pitch register, this process also involved a shift in ERB units (Greenwood, 1990) that centered the donor pitch contour around the mean of the receiver pitch contour. **Table 4** displays the mean (standard deviation), range, minimum and maximum standard deviation of the F0 per speaker group, separated by condition. The standard deviation is chosen as a measure of intonational variation because it gives an impression of the variability of speakers' pitch contours. It is used instead of pitch range (which also does this), because it is less sensitive to outliers (such as creaky voice, which has an extremely low F0).

After the manipulation step, the sound pressure level of each resynthesized utterance was normalized to 64 dB. Silent pauses in the utterances were excluded from this process. In total, the experiment consisted of 10 sentences × 4 speakers × 8 conditions = 320 utterances per task. The items were pseudorandomized with the criterion that at least seven other sentences had to separate two sentences with the same content. The resulting list of stimuli was then mirrored to create an additional list used to control for any learning effects.

Table 4 *Mean (SD), range, minimum and maximum standard deviation of the F0 in ERB per speaker group, separated by condition*

Condition	Speaker Group	Mean (SD)	Range	Minimum	Maximum
C1	L2_intermediate	1.24 (.19)	.75	.80	1.55
	L2_advanced	1.28 (.27)	1.14	.65	1.79
C2	L2_intermediate	1.87 (.45)	1.43	.99	2.42
	L2_advanced	1.72 (.41)	1.35	.94	2.30
C3	L2_intermediate	1.28 (.23)	.96	.78	1.74
	L2_advanced	1.35 (.29)	1.24	.60	1.85
C4	L2_intermediate	1.22 (.22)	.89	.75	1.64
	L2_advanced	1.28 (.28)	1.13	.65	1.78
C5	L2_intermediate	1.85 (.46)	1.44	.98	2.42
	L2_advanced	1.74 (.43)	1.37	.93	2.30
C6	L2_intermediate	1.78 (.47)	1.79	.77	2.56
	L2_advanced	1.67 (.40)	1.36	.89	2.26
C7	L2_intermediate	1.22 (.23)	.96	.77	1.73
	L2_advanced	1.28 (.31)	1.15	.66	1.81
C8	L2_intermediate	1.79 (.48)	1.76	.79	2.54
	L2_advanced	1.69 (.41)	1.36	.89	2.26
Donor	L1	1.58 (.42)	1.33	.84	2.17

5.4.4 Procedure

Even though the experiment was presented to the participants in an online format, experimental sessions took place in a quiet computer room, to ensure that all participants performed the experiment in equal conditions and to minimize distractions during the task. To that same end, the stimuli were presented to the participants through headphones. Sessions were performed in a group setting and in order to prevent unreliable answers due to fatigue or boredom, participants listened to the stimuli in three sets of roughly 10 minutes each, which was how long it generally took participants to rate one third of the stimuli. The stimuli blocks were separated by 10-minute breaks. During the breaks, the stimuli were blocked to the participants and communication concerning the stimuli or the task between the participants was not allowed. In total, the experiment took approximately 60 minutes. Qualtrics randomly assigned participants to either the comprehensibility task or the accentedness task. Participants were instructed to listen carefully to the utterances, as the differences between the stimuli would be subtle, and they were encouraged to use the complete range of the 7-point scale on which they were to judge either the accentedness or comprehensibility of the items. Before starting the rating task, the participants answered questions about their age, nationality, L1 and possible L2s, and the presence of any aural and/or visual impairments to ensure that they met the requirements explained above.

5.5 Results

First, the accentedness ratings were transformed, in such a way that the accentedness and comprehensibility ratings reflect the same direction of effect. Therefore, in the subsequent analyses higher ratings indicate higher comprehensibility and lower accentedness, in other words: more target-like perception. Subsequently, IBM SPSS Statistics 22.0.0 (IBM Corporation, 2017) was used to perform two repeated measures analyses of variance (ANOVAs) with Proficiency of the L2 speaker producing the receiver material (binary: intermediate or advanced) and Intonation, Rhythm, and Speech Rate (binary: either transferred from the donor to the receiver or not) within-subject variables, and the accentedness and comprehensibility ratings as dependent variables. Both the ANOVA for accentedness and comprehensibility revealed a main effect of Proficiency. For comprehensibility, a significant interaction was also found between Proficiency and Speech Rate, as well as between Proficiency, Speech Rate, and Intonation. These (interaction) effects of Proficiency proved to be straightforward in the sense that stimuli by less proficient Spanish learners of Dutch were systematically rated as more accented and more difficult to understand than stimuli by more proficient Spanish learners of Dutch. Thus, in order to facilitate the interpretation of our statistical analyses, we will report only the interactions between the three prosodic cues in this section. The complete overview of all main and interaction effects can be found in the **Appendix Tables A1 and A2**. For the mean accentedness and comprehensibility ratings per manipulation condition, see **Figure 2**.

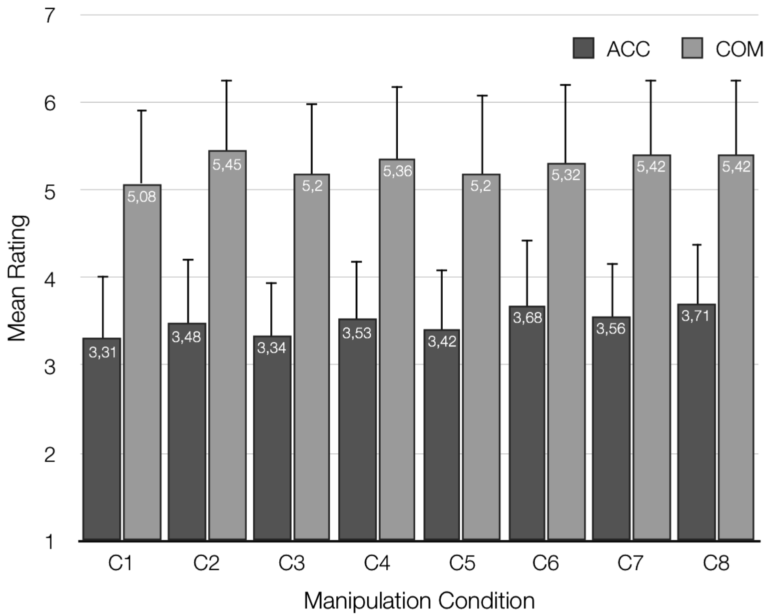


Figure 2. Mean accentedness and comprehensibility ratings per manipulation condition. Error bars represent the standard deviation.

5.5.1 Accentedness

All experimental conditions (i.e., C2-C8) received higher mean ratings than the control condition (i.e., C1), which indicates that using L1 donor intonation, rhythm, and speech rate all contribute to a lower perception of accentedness by L1 speakers, see **Table 5** for mean ratings and standard deviations per condition. Furthermore, all experimental conditions that included a transfer of one or a combination of the three prosodic features (i.e., C2-C7) received lower mean rating than the condition with the full manipulation (i.e., C8). This suggests that transferring a combination of intonation, rhythm and speech rate generates the most target-like perception in L1 listeners. As for the contribution of each specific prosodic feature to this perception, the analysis revealed a significant main effect of Intonation, $F(1, 29) = 54.31$, $p < .001$, $\eta_p^2 = .652$, as well as a significant main effect of Speech Rate, $F(1, 29) = 11.64$, $p = .002$, $\eta_p^2 = .286$. Rhythm was not found to significantly affect accentedness ratings, $F(1, 29) = .06$, $p = .806$. The analysis revealed no significant interactions between the three independent variables.

5.5.2 Comprehensibility

Once again, all experimental conditions (i.e., C2 – C8) received higher mean ratings than the control condition (i.e., C1), which indicates that using L1 donor intonation, rhythm and speech rate all contribute to a higher L1 perception of comprehensibility of speech by Spanish learners of Dutch, see **Table 5**. Furthermore, all experimental conditions that included a transfer of one or a combination of the three prosodic features (i.e., C3 - C7) received lower mean comprehensibility rating than the condition with the full manipulation (i.e., C8), with the exception of the condition in which only Intonation was manipulated (i.e., C2). This suggests that transferring a combination of intonation, rhythm and speech rate generates a favorable perception of comprehensibility in L1 listeners, and because the intonation-only condition received even higher ratings, it implies that this is at least partially due to the Intonation manipulation. Indeed, the analysis revealed a significant main effect of Intonation on the ratings for comprehensibility, $F(1, 29) = 11.39$, $p = .002$, $\eta_p^2 = .282$, as well as a significant, but smaller, main effect of Speech Rate, $F(1, 29) = 4.35$, $p = .046$, $\eta_p^2 = .130$. Rhythm was not found to affect comprehensibility ratings in a significant way, $F(1, 29) = .12$, $p = .727$. Furthermore, the analysis revealed significant interactions between Intonation and Rhythm, $F(1, 29) = 18.89$, $p < .001$, $\eta_p^2 = .394$, between Intonation and Speech Rate, $F(1, 29) = 27.19$, $p < .001$, $\eta_p^2 = .484$, and between Rhythm and Speech Rate, $F(1, 29) = 21.94$, $p < .001$, $\eta_p^2 = .431$. Finally, a three-way interaction was also found between all independent variables, $F(1, 29) = 17.73$, $p < .001$, $\eta_p^2 = .379$

Table 5 *Mean accentedness and comprehensibility ratings per manipulation condition*

Condition	Accentedness		Comprehensibility	
	Mean	*(SD)*	*Mean*	*(SD)*
C1	1.24	(.19)	.75	.80
C2	1.87	(.45)	1.43	.99
C3	1.28	(.23)	.96	.78
C4	1.22	(.22)	.89	.75
C5	1.85	(.46)	1.44	.98
C6	1.78	(.47)	1.79	.77
C7	1.22	(.23)	.96	.77
C8	1.79	(.48)	1.76	.79

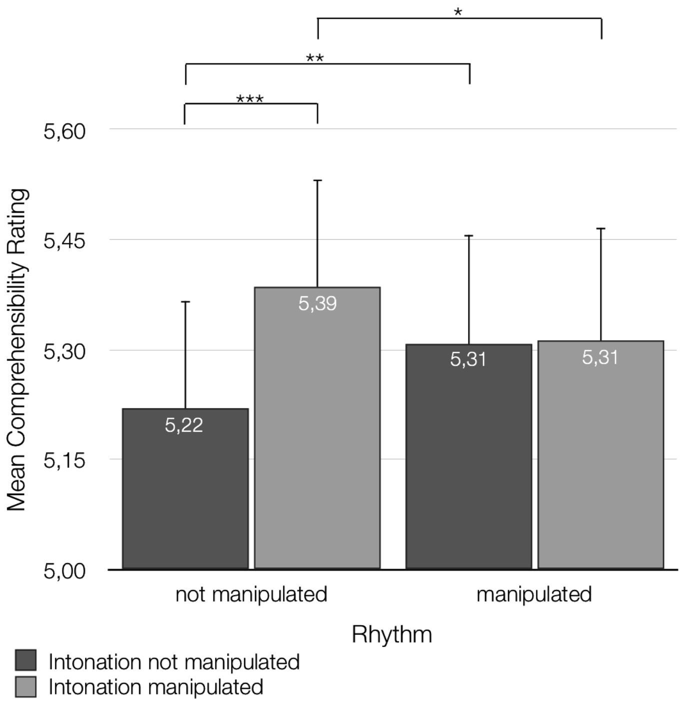


Figure 3. Mean comprehensibility ratings per Intonation condition, separated by Rhythm condition. Error bars represent the standard error. *Note:* * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

In order to further explore the interactions found between the independent variables, pairwise comparisons using the Bonferroni method were performed. The interaction between Rhythm and Intonation is illustrated in **Figure 3**, which shows that there is always a significant difference between the conditions in which Rhythm is manipulated (i.e., C3, C5, C7, and C8) and those in which it is not manipulated (i.e., C1, C2, C4, and C6), irrespective of whether Intonation is manipulated or not. Conversely, there is only a significant difference between the conditions in which Intonation is manipulated (i.e., C2, C5, C6, and C8) and those in which it is not manipulated (i.e., C1, C3, C4, and C7) when Rhythm is not manipulated. When Rhythm is manipulated, there is no significant difference between the two Intonation conditions.

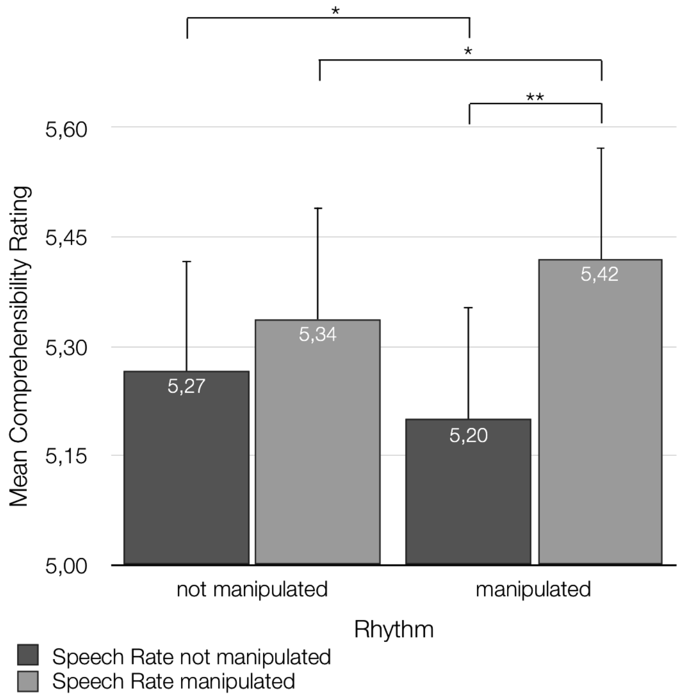


Figure 4. Mean comprehensibility ratings per Speech Rate condition, separated by Rhythm condition. Error bars represent the standard error. *Note:* * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Figure 4 illustrates the interaction between Speech Rate and Rhythm and shows that there is a significant difference between the conditions in which Rhythm is manipulated and those in which it is not manipulated, irrespective of whether Speech Rate is manipulated or not. The difference between the conditions in which Speech Rate is manipulated and those in which it is not manipulated, is only significant when Rhythm is manipulated. When Rhythm is not manipulated, there is no difference between the two Speech Rate conditions.

5

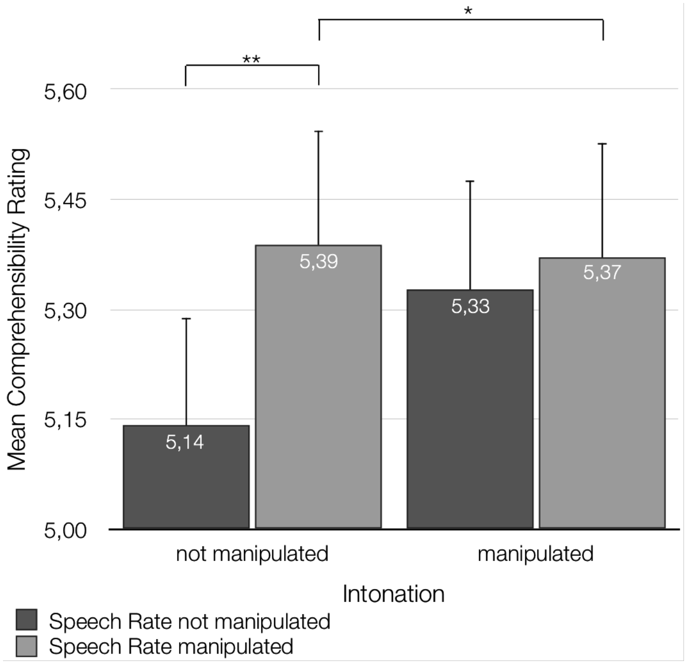


Figure 5. Mean comprehensibility ratings per Speech Rate condition, separated by Intonation condition. Error bars represent the standard error. *Note:* * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Concerning the interaction between Speech Rate and Intonation, **Figure 5** reveals that there is only a significant difference between the conditions in which Intonation is manipulated and those in which it is not manipulated, when Speech Rate is not manipulated. When Speech Rate is manipulated the difference is not significant. The analysis only reveals a significant difference between the conditions in which Speech Rate is manipulated and those in which it is not manipulated when Intonation is not manipulated. When Intonation is manipulated, there is no difference between the two Speech Rate conditions.

Finally, **Figure 6** illustrates the three-way interaction between the three prosodic features. This interaction is consistent with the found main effects, in that the manipulation of Rhythm only leads to significant changes in comprehensibility ratings in combination with the other two prosodic cues. It is also in line with the two-way interactions found between Intonation and Rhythm, as well as between Intonation and Speech Rate. The fact that there is a significant three-way interaction appears to be due to the fact that when Intonation is manipulated, the results of the interaction between Rhythm and Speech Rate remain similar to when they were compared directly, but when Intonation is not manipulated, there is always a significant difference between Speech Rate conditions, irrespective of whether Rhythm is manipulated or not. In sum, for comprehensibility, the interplay between the three prosodic features is quite complex, which confirms the relevance of studying these prosodic cues in combination.

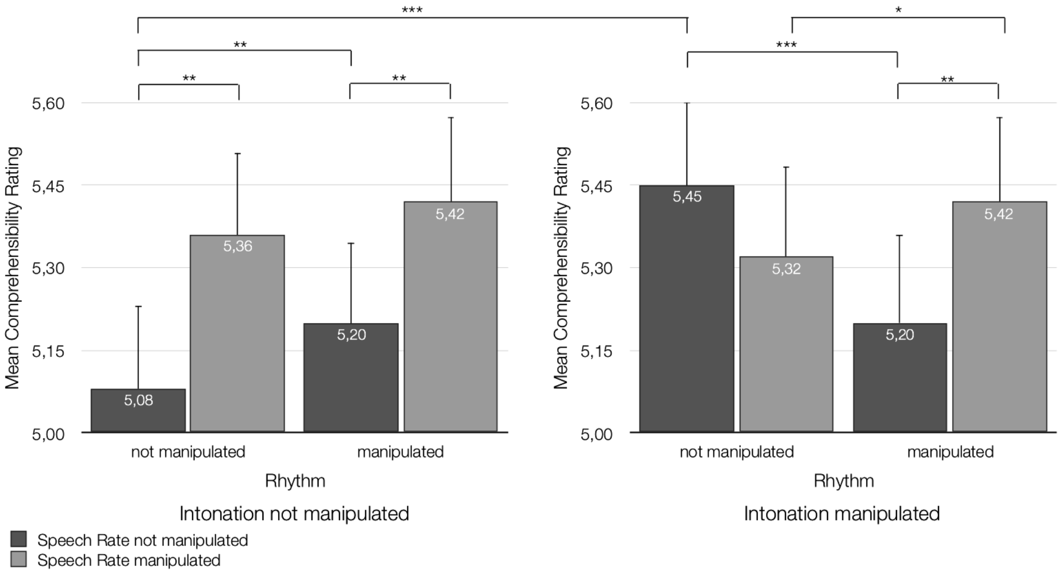


Figure 6. Mean comprehensibility ratings per Speech Rate condition, separated by Rhythm condition when Intonation is not manipulated (left) and when Intonation is manipulated (right). Error bars represent the standard error. *Note:* * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

5.6 Discussion and conclusion

The current study investigated whether transplanting the speech rate, intonation, and/or rhythm of a Dutch native speaker onto speech material by Spanish learners of Dutch contributed to L1 Dutch perceptions of accentedness and comprehensibility. In general, the results show that transplanting intonation, speech rate, and rhythmic properties positively affects L1 listeners' perceptions of foreign accent and ease of comprehension in all cases, but that not all improvements are significant. More specifically, accentedness ratings appear to only be influenced by the speakers' intonation and speech rate, and by both factors equally, while the contribution of the three independent variables to comprehensibility ratings is less straightforward. Whereas accentedness ratings are not influenced by the rhythm of the stimulus, comprehensibility ratings do appear to be affected by all three prosodic cues, albeit not to the same degree in all contexts: Intonation appears to contribute most to L1 perceptions of the comprehensibility of L2 speech, while speech rate also does so, but to a lesser extent. No main effect of rhythm on comprehensibility ratings was found, but rhythm in combination with intonation and/or speech rate did significantly affect comprehensibility ratings.

As main effects of speech rate (prediction 1) and intonation (prediction 2) have indeed been found for both accentedness and comprehensibility, our results are congruent with most of the previously mentioned studies (Magen, 1998; Munro & Derwing, 2001; Polyanskaya et al., 2016; Saito, Trofimovich & Isaacs, 2016; Van Els & De Bot, 1987; Chapter 3 of this dissertation). As mentioned in prediction 2, there were also two studies that did not find a significant effect of Intonation on accentedness, that is, Ramus and Mehler (1999) and Polyanskaya et al. (2016). Interestingly, these were also the two studies that reported a significant effect of Rhythm on accentedness (prediction 3), which is not replicated in the present study. In the case of Ramus and Mehler (1999), the discrepancy between their results and those of the current study might be due to the fact that their stimuli were delexicalized so that utterances became nonsensical: depending on the prosodic cue under manipulation, they transformed all segments into different variants of *sasasa,* replacing all consonants with /s/ and all vowels with /a/. Naturally, this process yields controlled stimuli that only differ in prosodic features, but it also results in stimuli that no longer contain any meaning that can serve as a context for discourse-dependent intonation. Conversely, the present study used utterances with meaning, which may serve as a semantic background for intonational-based features, such as prominence or finality. As intonation is highly context-dependent, it might have been difficult for the participants of Ramus and Mehler (1999) to distinguish between two languages based on this cue without any context at all, which may explain why they found no effect of intonation on accentedness. Alternatively, a repetition of similar syllables in their utterances may have a clearer impact on ratings of rhythm than is the case with utterances that are more variable in the make-up of syllable structures. In addition, the current study measures L1 perceptions of foreign accentedness, while Ramus and Mehler asked French natives to distinguish between English and Japanese utterances based on their rhythm, intonation and broad phonotactics. Their results are in line with other studies (e.g., Ramus et al., 2003), which report that listeners are able to distinguish between languages with very distinct rhythm (e.g., a Romance language such as Spanish vs. a Germanic language like Dutch), but often fail to discriminate languages with similar but not identical rhythm (e.g., two languages from the same language family, such as Dutch and English). It can be argued that assessing the foreign accentedness of speech based only on prosodic cues requires more sensitivity to minimal differences than distinguishing between two options that clearly differ in rhythm. Therefore, it is conceivable that the effect found by Ramus and Mehler (1999) simply was not strong enough to surface in the current study as well.

Regarding Polyanskaya et al. (2016), there are several differences between their study and the present research: first, Polyanskaya et al. base their rhythm analysis on rhythm metrics instead of lengthening measures. Hence, their rhythm manipulation takes place at the phonemic level, rather than the syllabic level, as is the case in the current study. Second, in their study, L2 speech rate and/or rhythm are transplanted onto L1 segmental (and in some conditions intonational) material, while in the current study the opposite is done. Our choice was based on the assumption that the presence of segmental errors is a typical characteristic of L2 speech, and as such, our stimuli would more closely resemble L2 speech as L1 listeners might encounter it in real life. Both of these differences might have contributed to the inconsistency in findings, as previous studies have shown that off-target production of segments (and thus the manipulation of rhythm through segments) might influence syllable complexity, which, in turn, affects rhythm (e.g., Payne, Post, Prieto, Vanrell & Astruc, 2012; Chapter 4 of this dissertation). The nature of the pitch manipulation might explain the contradictory findings with respect to the effect of intonation: While they used utterances with a flat F0, the current study relied on utterances with L2 intonation as a baseline condition. Arguably, the difference between a monotonized pitch contour and typical L1 intonation is less salient and obstructing than the difference between L2 and L1 intonation. Using actual L2 intonation requires the listener to mentally recognize and correct atypical intonation patterns, as L2 speech generally contains noticeable lexical and phrasal stress errors (e.g., Magen, 1998; Chapter 2 of this dissertation). Listening to speech with monotonized pitch intuitively takes a lower toll on cognitive load, as it may be more difficult to ignore an atypical pattern, than to not be confronted with said pattern in the first place.

In sum, our study has shown that both accentedness and comprehensibility are affected as a result of manipulations of intonation and speech rate, but not as a result of the manipulation of rhythm. In this sense, our study corroborates the idea that it is important to investigate several prosodic cues in the same design instead of treating prosody as one holistic cue: the three prosodic features manipulated in the current study affected L1 perceptions of accentedness and comprehensibility of L2 speech to different degrees, and only the interaction between the three cues significantly affected comprehensibility. In addition, both perceptual measures are influenced differently by the manipulation of the three prosodic features: While rhythm did interact significantly with both other factors, as well as their combination, in our comprehensibility analyses, no significant interactions with rhythm were found for accentedness. In addition, the effects of Intonation and Speech Rate on accentedness ratings were comparable, whereas the contribution of Intonation to comprehensibility was stronger that the contribution of Speech Rate. In this sense, our study corroborates previous studies suggesting that it is important to clearly differentiate between different aspects of perception in the field of L2 acquisition, as not all measures are affected equally. Therefore, future research might elaborate on the current study by also investigating the effect of these three prosodic cues on intelligibility. As such a study requires the collection of a substantial stimuli set containing enough different stimuli to enable reliable measurements of for instance reaction times, eye-tracking or transcription tasks, it was impossible to incorporate this measure in the current experiment, which, out of necessity, made use of ten utterances recorded for Chapter 4 of this dissertation.

Our study also has didactic consequences as it shows that improvements in L2 intonation and speech rate significantly influence L1 perceptions of both the accentedness and comprehensibility of L2 speech. While it is impossible to conclude whether these effects also hold for actual processing by L1 speakers for the moment, L2 learners might be motivated to facilitate communication with natives by reducing their perceptions of accentedness and difficulty of comprehension. Luckily for L2 learners, speech rate is known to increase with proficiency level (Derwing & Munro, 2001), requiring little extra coaching. In contrast, specific training to improve intonation has been proven to be effective (e.g., Ramírez Verdugo, 2006) and therefore may be a useful addition to the L2 curriculum. Further research might therefore be dedicated to the effect of different types of training, as well as the use of training in the different functions of intonation and the contexts in which they can be employed.

The finding that intonation and speech rate affect comprehensibility and accentedness may also be useful within the fields of speech pathology and automatic speech recognition. As is the case with L2 teachers, speech pathologists may choose to dedicate specific attention to intonation training in order to improve the ease with which interlocutors understand their clients. And while the application of specific intonational patterns remains a challenge in automatic speech recognition and text-to-speech software due to its context-dependent nature, it is worth reiterating that this element does appear to be essential to L1 perceptions of foreign accentedness and ease of comprehension of L2 speech.

Notes

---

1 Winters and O'Brian's (2013) manipulation of syllable durations overlaps with both our rhythm and speech rate manipulations as the syllable durations in their study were converted into the durations of the syllables produced for the same sentence in the other L1 talker group. This means that not only the relative syllable durations (e.g., of prominent syllables relative to non-prominent syllables) have been manipulated (related to the concept of rhythm), but also the absolute syllable durations (related to changes in speech rate).

2 Rhythm arguably incorporates more than just timing differences. In the current study, this is taken into consideration as our rhythm manipulation reflects the accentual and final lengthening patterns that are typical of Dutch but not of Spanish. Therefore, our manipulation reflects the durational variation that exists in Dutch as a result of the boundary and prominence marking, without it being intertwined with intonation, as it has been in previous studies.

3 While it would have been interesting to include a control condition with both the segmental string and the prosodic features of an L1 speaker of Dutch, we chose to limit the length of the experiment by only including a control condition in which both the segmental string and the prosodic features were produced by a Spanish learner of Dutch as this condition would most resemble L2 Dutch as it is heard by Dutch natives, segmental mistakes usually being quite salient in L2 speech.

References

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, *42*(4), 529-555.

Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning*, *38*(4), 561-613.

Beckman, M., Díaz-Campos, M., McGory, J. T., & Morgan, T. A. (2002). Intonation across Spanish, in the Tones and Break Indices framework. *Probus*, *14*(1), 9-36.

Boersma, P., & Weenink, D. (2010). Praat: Doing phonetics by computer (Version 6.0.21) [Computer software]. Retrieved from http://www.fon.hum.uva.nl/praat/

Bolinger, D. L. (1965). Pitch accent and sentence rhythm. In I. Abe & T. Kanekiyo (Eds.), *Forms of English: Accent, morpheme, order* (pp. 139–180). Cambridge, MA: Harvard University Press.

Booij, G. (1995). *The phonology of Dutch*. Oxford, United Kingdom: Clarendon Press.

Boula de Mareüil, P., & Vieru-Dimulescu, B. (2006). The contribution of prosody to the perception of foreign accent. *Phonetica*, *63*(4), 247-267.

Cambier-Langeveld, T., Nespor, M., & van Heuven, V. J. (1997). The domain of final lengthening in production and perception in Dutch. *Proceedings of the European Conference on Speech Communication and Technology*, *2*, 931-934.

Cambier-Langeveld, T., & Turk, A. E. (1999). A cross-linguistic study of accentual lengthening: Dutch vs. English. *Journal of Phonetics*, *27*(3), 255-280.

Caspers, J., & Horłoza, K. (2012). Intelligibility of non-natively produced Dutch words: Interaction between segmental and suprasegmental errors. *Phonetica*, *69*(1-2), 94-107.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge, United Kingdom: Cambridge University Press.

De Johnson, T. H., O'Connell, D. C., & Sabin, E. J. (1979). Temporal analysis of English and Spanish narratives. *Bulletin of the Psychonomic Society*, *13*(6), 347-350.

Delattre, P. (1966). A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics in Language Teaching*, *4*, 183–198.

Delattre, P. (1969). An acoustic and articulatory study of vowel reduction in four languages. *International Review of Applied Linguistics for Language Teaching, 7*, 295-325.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in Second Language Acquisition*, *19*(1), 1-16.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quaterly*, *39*(3), 379-397.

Edmunds, P. (2009). ESL speakers' production of English lexical stress: The effect of variation in acoustic correlates on perceived intelligibility and nativeness (Doctoral dissertation). The University of New Mexico, United States.

Guion, S. G., Flege, J. E., Liu, S. H., & Yeni-Komshian, G. H. (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics*, *21*(2), 205-228.

Gussenhoven, C. (2005). Transcription of Dutch Intonation. In: Jun, Sun-Ah (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing* (pp. 118-145). Oxford, United Kingdom: Oxford University Press.

Greenwood, D. D. (1990). A cochlear frequency-position function for several species—29 years later. *The Journal of the Acoustical Society of America*, *87*(6), 2592-2605.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, *38*(2), 201-223.

Hartsuiker, R. J. (2002). The addition bias in Dutch and Spanish phonological speech errors: The role of structural context. *Language and Cognitive Processes*, *17*(1), 61-96.

Hirst, D., & Di Cristo, A. (Eds.). (1998). *Intonation systems: a survey of twenty languages*. Cambridge University Press.

Hualde, J. I. (2005). *The Sounds of Spanish with Audio CD*. Cambridge, United Kingdom: Cambridge University Press.

IBM Corporation. (2017). IBM SPSS Statistics (version 22.0.0). [Computer software]. Retrieved from http://www.spss.com/software/statistics.

Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, *58*(2), 541-573.

Jilka, M. (2000). The contribution of intonation to the perception of foreign accent: Identifying intonational deviations by means of F0 generation and resynthesis (Doctoral dissertation). Universität Stuttgart, Stuttgart, Germany.

Koopmans-Van Beinum, F. J. (1980). *Vowel contrast reduction: An acoustic and perceptual study of Dutch vowels in various speech conditions*. Amsterdam: Academische Pers.

Li, A., & Post, B. (2014). L2 acquisition of prosodic properties of speech rhythm. *Studies in Second Language Acquisition*, *36*(2), 223-255.

Magen, H. S. (1998). The perception of foreign-accented speech. *Journal of Phonetics*, *26*(4), 381-400.

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*, *32*(4), 543-563.

Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication*, *9*(5-6), 453-467.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *49*(s1), 285-310.

Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech the role of speaking rate. *Studies in Second Language Acquisition*, *23*(4), 451-468.

Navarro Tomás, T. (1966). *Estudios de fonologia española* (Studies on Spanish phonology, in Spanish). New York: Las Américas.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human perception and performance*, *24*(3), 756-766.

Payne, E., Post, B., Prieto, P., Vanrell, M., & Astruc, L. (2012). Measuring child rhythm. *Language and Speech*, *55*, 203–229.

Pettorino, M., & Vitale, M. (2012). Transplanting native prosody into second language speech. In M. G. Busà, and A. Stella (eds.), *Methodological Perspectives on Second Language Prosody: Papers from ML2P 2012* (pp. 11-16). Padova: CLEUP.

Pike, K. L. (1945). *The intonation of American English*. Ann Arbor: University of Michigan.

Polyanskaya, L., Ordin, M., & Busa, M. G. (2016). Relative Salience of Speech Rhythm and Speech Rate on Perceived Foreign Accent in a Second Language. *Language and Speech, 60(3), 333-355*.

Prieto, P., Vanrell, M., Astruc, L., Payne, E., & Post, B. (2012). Phonotactic and phrasal properties of speech rhythm: Evidence from Catalan, English, and Spanish. *Speech Communication*, *54*, 681–702.

Qualtrics. (2016) [Computer software]. Retrieved from http://www.qualtrics.com/

Quené, H., & Van Delft, L. E. (2010). Non-native durational patterns decrease speech intelligibility. *Speech Communication*, *52*(11), 911-918.

Ramírez Verdugo, M. D. (2003). Non-native interlanguage intonation systems: A study based on a computerized corpus of Spanish learners of English. *ICAME Journal*, *26*, 115-132.

Ramírez Verdugo, M. D. (2006). A study of intonation awareness and learning in non-native speakers of English. *Language Awareness*, *15*(3), 141-159.

Ramus, F., Dupoux, E., & Mehler, J. (2003). The psychological reality of rhythm classes: Perceptual studies. In M. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 337–342). Barcelona: Universitat Autonomá de Barcelona.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *The Journal of the Acoustical Society of America*, *105*(1), 512-521.

Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, *37*(2), 217-240.

Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics*, *25*(1), 1-24.

Van Els, T., & De Bot, K. (1987). The role of intonation in foreign accent. *The Modern Language Journal*, *71*(2), 147-155.

Verhoeven, J., De Pauw, G., & Kloots, H. (2004). Speech rate in a pluricentric language: A comparison between Dutch in Belgium and the Netherlands. *Language and Speech*, *47*(3), 297-308.

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, *35*, 501–522.

Winters, S., & O'Brien, M. G. (2013). Perceived accentedness and intelligibility: The relative contributions of F0 and duration. *Speech Communication*, *55*(3), 486-507.

Yoon, K. (2007). Imposing native speakers' prosody on non-native speakers' utterances: The technique of cloning prosody. *Journal of the Modern British and American Language & Literature*, *25*(4), 197-215.

Appendix

Table A1 *Accentedness: overview of relevant main effects and interactions*

Independent variables(s)	Effect
Proficiency	**$F(1, 29) = 34.484, p < .001, \eta_p^2 = .543$**
Intonation	**$F(1, 29) = 54.305, p < .001, \eta_p^2 = .652$**
Rhythm	$F(1, 29) = .061, p = .806, \eta_p^2 = .002$
Speech Rate	**$F(1, 29) = 11.642, p = .002, \eta_p^2 = .286$**
Proficiency*Intonation	$F(1, 29) =. 127, p = .724, \eta_p^2 = .004$
Proficiency*Rhythm	$F(1, 29) = .126, p = .726, \eta_p^2 = .004$
Proficiency*Speech Rate	$F(1, 29) = .009, p = .925, \eta_p^2 = .000$
Intonation*Rhythm	$F(1, 29) = .980, p = .330, \eta_p^2 = .033$
Intonation*Speech Rate	$F(1, 29) = .841, p = .367, \eta_p^2 = .028$
Rhythm*Speech Rate	$F(1, 29) = 1.551, p = .223, \eta_p^2 = .051$
Intonation*Rhythm*Speech Rate	$F(1, 29) = 1.624, p = .213, \eta_p^2 = .053$
Proficiency*Intonation*Rhythm*Speech Rate	$F(1, 29) = .561, p = .460, \eta_p^2 = .019$

Note. Bold characters are used to mark significant effects.

Table A2 *Comprehensibility: overview of relevant main effects and interactions*

Independent variables(s)	Effect
Proficiency	**$F(1, 29) = 86.382, p < .001, \eta_p^2 = .702$**
Intonation	**$F(1, 29) = 11.388, p = .002, \eta_p^2 = .282$**
Rhythm	$F(1, 29) = .124, p = .727, \eta_p^2 = .004$
Speech Rate	**$F(1, 29) = 4.347, p = .046, \eta_p^2 = .130$**
Proficiency*Intonation	$F(1, 29) = .055, p = .817, \eta_p^2 = .002$
Proficiency*Rhythm	$F(1, 29) = .000, p = .990, \eta_p^2 = .000$
Proficiency*Speech Rate	**$F(1, 29) = 7.908, p = .009, \eta_p^2 = .214$**
Intonation*Rhythm	**$F(1, 29) = 18.892, p < .001, \eta_p^2 = .394$**
Intonation*Speech Rate	**$F(1, 29) = 27.185, p < .001, \eta_p^2 = .484$**
Rhythm*Speech Rate	**$F(1, 29) = 21.938, p < .001, \eta_p^2 = .431$**
Intonation*Rhythm*Speech Rate	**$F(1, 29) = 17.731, p < .001, \eta_p^2 = .379$**
Proficiency* Intonation*Rhythm*Speech Rate	$F(1, 29) = .405, p = .529, \eta_p^2 = .014$

Note. Bold characters are used to mark significant effects.

5

6

6

General Discussion & Conclusion

This dissertation focused on the production and perception of intonation and rhythm by native (L1) and non-native (L2) speakers of Spanish and Dutch. In order to investigate which factors are relevant to successful prosody acquisition, as well as which prosodic features contribute most to L1 perceptions of accentedness, comprehensibility, and intelligibility, we performed four empirical studies. Taken together, these studies were set up to answer the broad research questions that we formulated in Chapter 1 of this dissertation. In this final chapter, we discuss to what extent our four studies have provided answers to these research questions and reflect on a number of theoretical and methodological issues that span more than one chapter, as well as the practical implications of our findings. We end this chapter with a discussion of the limitations of our work, suggestions for future research, and some concluding remarks.

6.1 Main research questions and summary

The studies reported in this dissertation all focused on the production of L2 prosody in speech by L1 and/or L2 speakers of Spanish and Dutch or the perception thereof by L1 speakers of Dutch. As each study was specifically set up to find the answer to one of the main research questions, we will repeat each question here, before turning to its answer.

6.1.1 L2 intonation production

RQ1 Do the differences between Dutch and Spanish in the way they use intonation to mark focus lead to transfer effects in the L2 speech of learners of these languages? If so, does proficiency level influence this effect?

In **Chapter 2** of this dissertation, we reported an experiment on the production of intonation in the L1 and the L2 by Dutch and Spanish natives, as well as Dutch learners of Spanish (DLS) and Spanish learners of Dutch (SLD) of varying proficiency levels. We elicited semi-spontaneous speech by asking participants to describe objects and their colors (e.g., a red mitten) in varying focal conditions. In the first condition, both elements of the target NP were new in the context; in the second, only the second element was new, while the first one was given within the context; in the third, the first element was new, while the second was given; and in the fourth, both elements were given in the context. By comparing the pitch accent distributions in these four conditions for all language groups, we demonstrated that, as predicted, L1 Dutch and L1 Spanish differed in the way they use pitch accents to mark focus, which is in line with previous work on these two languages, which had not been compared directly on this prosodic property before (Face, 2002; Hualde, 2005; Krahmer & Swerts, 2001; Nooteboom & Kruyt, 1987; Swerts, Krahmer & Avesani, 2002). Moreover, we found that these differences between L1 Dutch and L1 Spanish indeed led to transfer effects in the L2 speech of DLS and SLD. The latter produced a pitch accent distribution in which the final element was accented in all focal contexts, irrespective of its focal status, while the speech of the former was characterized by the accentuation of focused elements, irrespective of their position in the noun phrase. As these pitch accent distributions are typical of their L1, but not of their L2 respectively, we concluded that prosodic transfer indeed took place in speech by both DLS and SLD. In this sense, our findings were in line with previous studies on L2 acquisition of intonation to mark focus that relied on subjective prominence judgments, instead of objective acoustic measurements (e.g., Rasier & Hiligsmann, 2009; Swerts & Zerbian, 2010).

In addition, our study showed that the extent to which L2 learners successfully produced target-like pitch accent distributions in the L2 was modulated by their L2 proficiency level: both DLS and SLD increasingly approached their target as their proficiency level increased. This is congruent with prior studies on the prosodic marking of focus structure in the L2 for other language combinations but had never been investigated for Dutch and Spanish before (e.g., Reichle & Birdsong, 2014; Swerts & Zerbian, 2010; Zubizarreta & Nava, 2011). Interestingly, this study not only enabled us to investigate transfer effects from the L1 to the L2, but also in the opposite transfer direction, i.e., from the L2 to the L1. With respect to the latter, our results showed that as L2 learners became more proficient, their L2 started to influence their L1. This corroborates the only other study that we know of to date showing that there is more than just one possible prosodic transfer direction (Mennen, 2004).

6.1.2 L2 intonation perception

RQ2 How does deviance in the production of intonation to mark focus by L2 learners influence L1 listeners' perceptions of their speech concerning its non-nativeness, the ease with which it is understood, and actual processing?

Chapter 3 of this dissertation focused on the consequences of the previously found prosodic transfer effects from the L1 to the L2 for communication between L1 and L2 speakers of Dutch. Specifically, we performed a series of experiments to assess the effect of deviance in one specific feature, that is, the prosodic marking of focus, on several types of L1 perceptions (i.e., accentedness, nativeness, comprehensibility, and intelligibility), which are usually not combined in one design. It is important to distinguish between these different perceptual measures, because while the subjective ones, such as accentedness and comprehensibility, provide relevant information on the experience of the L1 listener, they do not necessarily reflect actual understanding or processing (i.e., intelligibility), which is essential for successful communication. Moreover, while previous work generally investigated intonational deviance in general (e.g., Derwing and Munro, 1997; Munro, 1995; Munro and Derwing, 1995; Munro and Derwing, 1999; Van Els and De Bot, 1987), our study took into account the communicative function (in this case, the marking of focus) of a specific pitch movement (here, pitch accent distributions within a noun phrase) and determined whether prosodic deviance, even if only represented by such a relatively subtle cue, could influence L1 perceptions on several levels.

The first experiment used unedited fragments collected for Chapter 2 of this dissertation produced by L1 speakers of Dutch, more proficient SLD, and less proficient SLD. Consequently, these speech samples did not only differ in their pitch accent distributions, but also in segmental accuracy, and possibly other suprasegmental properties. The results of this experiment showed that L1 Dutch listeners rated speech by proficient SLD as significantly more foreign-accented and less native-like than speech produced by L1 Dutch. Speech by less proficient SLD was, in turn, rated as more foreign-accented and less native-like than speech by proficient SLD. A similar pattern was observed with respect to the perceived ease of understanding (though within a smaller range, in line with Derwing & Munro, 1997; Munro & Derwing, 1999): L1 Dutch listeners considered speech increasingly more difficult to understand as the proficiency level of the speaker decreased.

The second experiment reported in this study controlled for additional segmental and suprasegmental deviance, as it asked L1 Dutch listeners to indicate which audio fragment was most native-like within a specific focus context; the one in which the elicitation context was congruent with the focus condition or the one that was actually elicited in the opposite focus condition. The experiment revealed that L1 Dutch listeners chose randomly between both options when presented with speech by less proficient SLD. This is in line with our expectations, as less proficient SLD produce the same pitch accent distribution in all focus conditions, which means that L1 Dutch listeners could not rely on pitch accent placement to infer the focal status of an element. Dutch natives performed better when listening to speech by more proficient SLD and were most successful when presented with utterances produced by L1 speakers of Dutch, in which the differences between typical and atypical pitch accent distributions were most salient. This shows that Dutch natives can distinguish between utterances based on pitch accent placement alone and can make the corresponding nativeness judgments. Taken together, the first two experiments of this study showed that L1 Dutch listeners have clear judgments about the accentedness, nativeness, and comprehensibility of both L1 and L2 speech and rely on the prosodic properties of the utterance to make these judgments.

The final experiment of this study was set up to determine whether the aforementioned judgments by L1 Dutch listeners also corresponded to objective measures of processing. Our intelligibility task, a reaction-times experiment in which Dutch natives indicated as quickly as possible whether the content of the utterance they heard matched the picture they saw on the computer screen, revealed that reaction times did not differ significantly between the conditions with typical and atypical pitch accent distributions. Even when confronted with speech by L1 speakers of Dutch, in which the difference between typical and atypical pitch accent placement was most salient, listeners did not respond significantly faster to the utterances with typical pitch accent distributions than to utterances with atypical pitch accent distributions. In this sense, our findings contradicted those of prior studies (e.g., Anderson-Hsieh, Johnson & Koehler, 1992; Hahn, 2004; Ito & Speer, 2008; Magen, 1998; Swerts & Vroomen, 2015; Van Heuven, Kruyt & De Vries, 1981). We tentatively attribute this to the nature of our stimuli, which consisted of semi-spontaneous speech, in which the pitch accent distributions can be less salient than in the read aloud speech stimuli used by most prior studies (De Ruiter, 2015). Hence, even though prosodic deviance in the distribution of pitch accents significantly affected all subjective measures of L1 perception (i.e., accentedness, nativeness, and comprehensibility), the only objective measure of L1 perception (i.e., intelligibility) remained unaffected, which is congruent with Munro and Derwing (1999), who used a holistic measure of intonation.

6.1.3 L2 rhythm production

RQ3 Does the direction in which L2 acquisition takes place affect the successful attainment of speech rhythm by Spanish learners of Dutch and Dutch learners of Spanish?

Chapter 4 of this dissertation was dedicated to L2 speech rhythm production and compared the accentual and final lengthening patterns in speech of L1 speakers of Dutch and Spanish, as well as DLS and SLD of varying proficiency levels. Based on previous work (e.g., Li & Post, 2014; White & Mattys, 2007), including the study presented in the second chapter of this dissertation, we expected that prosodic transfer would occur in the rhythm acquisition process and aimed to find out whether the degree of transfer would be influenced by the direction in which acquisition takes place. In other words, while Chapter 2 investigated *transfer* direction (i.e., from the L1 to the L2 and vice versa), Chapter 4 focused on the effect of *learning* direction (i.e., from language X, here Dutch, to language Y, here Spanish, and vice versa) on L2 prosody production. Based on the general acquisition pattern seen in L1 rhythm acquisition (e.g., Allen & Hawkins, 1978; Bunta & Ingram, 2007; Grabe, Watson & Post, 1999; Schmidt & Post, 2015), the syllable complexity constraints of Spanish and Dutch, and the degree of accentual and final lengthening used in both languages, we predicted that Dutch speech rhythm would be more difficult to acquire for SLD than Spanish speech rhythm would be for DLS. We based ourselves on the Markedness Differential Hypothesis (MDH), which states that "those areas of the target language which differ from the native language and are more marked that the native language will be difficult", while "those areas of the target language which are different from the native language but are not more marked than the native language will not be difficult." (Eckman, 1977, p. 321).

Our results showed that the target L2 rhythm was easier to acquire for DLS than for SLD, suggesting that learning direction is indeed an additional factor that plays a role in the L2 acquisition process. Thus, our results corroborated the MDH and the only two studies to our knowledge that applied the MDH to L2 prosody acquisition: Ordin and Polyanskaya (2015), who investigated L2 rhythm acquisition, but whose design did not allow for explicit claims with respect to the effect of learning direction, and Rasier and Hiligsmann (2009), who investigated L2 acquisition of pitch accents to mark focus by native speakers of Dutch learning French and French learners of Dutch.[1] Interestingly, a clear effect of syllable structure on the target-like production of L2 speech rhythm was also found in the sense that L2 learners performed systematically more on target when they produced sentences in which syllable structure was kept relatively simple than when they produced sentences in which relatively more complex syllables structures were used. This effect was more noticeable for SLD than for DLS. This led us to believe that the L2 rhythm acquisition process is characterized by an interplay between several phonetic properties, including phonotactic constraints and constraints on consonant cluster production (as in L1 rhythm acquisition, see Payne, Post, Prieto, Vanrell & Astruc, 2012). Due to these findings, we emphasized the need for L2 acquisition models that allow for the analysis of speech rhythm acquisition as a multi-systemic process (see 6.2.1).

6.1.4 L2 rhythm perception

RQ4 What is the relative contribution of intonation, rhythm and speech rate in speech produced by L2 learners on L1 listeners' perceptions of non-nativeness and ease of understanding?

The last empirical study of this dissertation (**Chapter 5**) concerned the perception of L2 rhythm by L1 Dutch listeners, but in contrast to most previous studies, it was not limited to the contribution of only one prosodic feature: L2 intonation and speech rate were also included in the design, in order to determine the relative contribution of L2 speech rate, rhythm, and intonation in speech by SLD to perceptions of accentedness and comprehensibility by L1 Dutch. The combination of several prosodic features in one design is uncommon, but relevant, as it enables us to investigate how these prosodic properties, that are normally simultaneously present in speech, interact with each other with respect to their effect on accentedness and comprehensibility. Using speech transplantation and resynthesis techniques, we manipulated the speech material of two less proficient and two more proficient SLD that participated in the study reported in Chapter 4: using their segmental strings as receiver material, we transplanted all possible combinations of the aforementioned prosodic properties taken from a L1 Dutch donor and resynthesized the resulting utterances. Dutch natives listened to the thus created utterances and rated them on the perceived foreign accent of the speaker and the perceived ease of comprehension of the utterance.

The results of this experiment showed that accentedness and comprehensibility were affected differently by deviance in (a combination of) intonation, speech rate, and rhythm, which is in line with our findings in Chapter 3 in which we exclusively manipulated intonation. We found that accentedness was only affected by the individual manipulations of intonation and speech rate, corroborating the findings of previous studies on the effect of those prosodic features in isolation (e.g., Magen, 1998; Munro & Derwing, 2001; Saito, Trofimovich & Isaacs, 2016; Van Els & De Bot, 1987; and Chapter 3 of this dissertation). Conversely, comprehensibility was affected by those two individual manipulations, but also by several interactions between those two, and even all three, prosodic cues. No main effect of rhythm was found on either accentedness or comprehensibility (which is incongruent with prior work by Ramus & Mehler, 1999, and Polyanskaya, Ordin & Busà, 2016, for which we offered several explanations). In sum, the results of this study emphasized that different L1 perception measures are affected in varying ways by prosodic deviance. In addition, it demonstrated the relevance of combining several prosodic factors into one design, as our study revealed that different prosodic features interacted with each other, and as such it more closely approached the type of interplay between these factors that would occur in natural L2 speech.

6.2 Overarching findings

Each empirical chapter of this dissertation contains its own discussion section highlighting the theoretical and methodological implications of that individual study, but when taken together there are a number of theoretical and methodological issues that span more than one chapter, which we discuss below.

6.2.1 The interplay between L2 features

Theoretically, all four studies reported in this dissertation, and especially in Chapters 4 and 5, supported the notion that the different prosodic features that we often distinguish between, e.g., intonation, rhythm, vowel quality, are actually interrelated. As mentioned in the General Introduction of this dissertation, prior work has shown that all four features generally used to measure prosody, i.e., intensity, pitch, length, and timbre, are used in the production of phrasal emphasis, also called phrasal intonation in Dutch (Sluijter, 1995). As we demonstrated in our rhythm production study (Chapter 4), phrasal intonation, in turn, affected measures of accentual lengthening, which are known to be congruent with our perception of speech rhythm (Prieto, Vanrell, Astruc, Payne & Post, 2012). Our last perception study (Chapter 5) also supported the idea of interplay between prosodic cues, as it showed that manipulations of intonation, speech rate, and rhythm interacted to affect L1 perceptions of comprehensibility. To further add to the complexity of the L2 prosody acquisition process, Chapter 4 showed that not only suprasegmental factors interacted in L2 prosody acquisition, but that segmental accuracy may also influence this process (cf. Ulbrich, 2015). In sum, our studies showed the need for a model of L2 (prosody) acquisition that accounts for the interplay between these different phonetic areas. In addition, such a model would ideally also account for the extra-linguistic factors that, as we have shown, affect the acquisition of L2 prosody, such as learning direction and the proficiency level of the L2 learner. Of course, there are many additional linguistic and extra-linguistic factors that are assumed to influence this process and that should be incorporated in a model of L2 acquisition, such as learner motivation, L2 aptitude and attitude, acquisition environment, the used learning method, the type and amount of input the L2 learner receives, and the amount of output they produce.

One model that claims to be able to integrate all such social and cognitive aspects of L2 acquisition, and in addition suggests that it is exactly the interplay between them that fuels L2 development, is the Dynamic Systems Theory (De Bot, Lowie & Verspoor, 2007). Where some L2 acquisition theories are difficult to apply to prosodic features (e.g., Best's Perceptual Assimilation Model, 1995, and Flege's Second Language Perception model, 1995), or are only compatible with certain theoretical frameworks (e.g., Archibald's formal model of the acquisition of L2 lexical stress, 1994, or Özçelik's Prosodic Acquisition Path Hypothesis, 2016), Dynamic Systems Theory is compatible with various broader theories about L2 acquisition (e.g., both Universal Grammar and Emergentism), as well as more specific hypotheses made with respect to L2 acquisition (e.g., Eckman's Markedness Differential

Hypothesis, 1977). Considering the fact that, to our knowledge, no study has ever investigated the acquisition of L2 prosody from a Dynamic Systems Theory perspective, this leaves an interesting opportunity for future work. In a similar vein, it would also be intriguing to explore how our results could be integrated into more recent models explicitly targeting L2 intonation acquisition, such as the L2 Intonation Learning theory (LILt, Mennen, 2015), and whether these would also generalize to the acquisition of L2 rhythm. This, too, is left as an exercise for future research.

6.2.2 The operationalization of prosodic atypicality and L1 perception

The studies reported in Chapter 3 and Chapter 5 showed that it is relevant to clearly define and distinguish between different measures of perception. While our intonation perception study in Chapter 3 revealed that accentedness and nativeness appear to represent the extremes of one and the same scale (in line with Edmunds, 2009; Schairer, 1992), it was also demonstrated that accentedness, comprehensibility, and intelligibility were affected differently and to different degrees by deviance in the production of pitch accent distributions to mark focus. This was corroborated by the results of Chapter 5, which also showed that accentedness and comprehensibility were not influenced similarly by prosodic deviance. The fact that intelligibility was not affected by prosodic deviance, while accentedness and comprehensibility were, might be explained by the fact that L1 listeners are not only extremely sensitive to phonetic errors and irregularities (Flege, 1984), but also accustomed to hearing and mentally correcting them. Even in the L1, listeners are continuously confronted with many types of phonetic variation, for example due to the gender, age, and regional background of their interlocutor, and cannot afford to let that variation impede communication. Hence, listeners are trained from an early age to compensate for such deviations from the norm, irrespective of whether they are produced by L1 or L2 speakers.

More importantly, the two studies also showed that combining deviance in various prosodic properties may change the relationship between different perception measures: while we reported similar rating patterns for accentedness and comprehensibility that only differed in their strength in Chapter 3, the findings of Chapter 5 showed that accentedness and comprehensibility may also be affected in different ways, irrespective of the strength of the effect, when deviance in several prosodic cues was combined. In sum, it appears risky to take the notions of 'perception' and 'prosodic deviance' too broadly, and future work might aim to use more fine-grained operationalizations of prosodic deviance, as well as measures of perception, in order to make studies within the field optimally comparable.

6.2.3 Spontaneous versus non-spontaneous stimuli

Another recurrent methodological issue in the studies reported in this dissertation is the type of speech used for stimuli in perception studies. We have used relatively more spontaneous and relatively less spontaneous speech stimuli in Chapter 3 and Chapter 5 of this dissertation respectively, and feel that, especially when investigating a specific prosodic cue in a specific communicative context (as is the case in Chapter 3, where we investigated the use of pitch accent distributions to mark focus) it is important to be aware of the effect speech style may have on the results of such a study.

While read-aloud speech may be perfectly suitable for studies assessing, for example, overall foreign-accentedness (where it indeed is common practice, e.g., Munro & Derwing, 2001), a more detailed measure might require an operationalization of the prosodic cue under investigation that more closely resembles how listeners would come across this prosodic feature in real life. This might be the case with pitch accents to mark focus, for which it has been shown that they are realized quite differently in read speech and spontaneous speech respectively (De Ruiter, 2015; Howell & Kadi-Hanifi, 1991; Laan, 1997; Swerts, Strangert, & Heldner, 1996). While we are aware that the collection of spontaneous speech stimuli can be a complicated matter, we do want to point out its relevance, as it may explain why our findings in Chapter 3 differ from those of some prior studies (e.g., Swerts & Vroomen, 2015).

6.3 Practical implications

Unfortunately, investigating whether formal training in speech rate, intonation, and/or rhythm in the L2 classroom actually improves L2 learners' performance with respect to measures of accentedness, comprehensibility, and intelligibility was outside the scope of this dissertation. However, we do feel that our studies yielded three practical implications that are worth emphasizing; three take-home messages for L2 teachers, if you will.

(1) Learning the prosody of a L2 is difficult, but possible

Our production studies showed that achieving native-like production in the production of L2 intonation and rhythm is generally very difficult for both SLD and DLS. While it is not impossible, there was substantial variation between the different utterances produced by a specific L2 leaner, as well as between L2 learners in general. Undoubtedly, this variation is due to (the interplay between) a myriad of factors, some of which we investigated in this dissertation (e.g., learning direction, general L2 proficiency, syllable complexity) and some of which we were unable to control for (e.g., musical training, language aptitude, L2 attitude, motivation). Practically, we feel it is encouraging for L2 teachers and L2 learners to note that improvement is possible and will co-occur with general improvements in proficiency. This especially holds for speech rate (Derwing & Munro, 2001), but our studies showed gradual improvement in pitch accent production (Chapter 2) and rhythm production (Chapter 4) as well.

(2) Training may improve L2 prosody acquisition

Our study on L2 speech rhythm production by DLS and SLD suggested that constraints on consonant cluster production, as well as on syllable complexity, affect L2 rhythm acquisition. While syllable structure may be difficult to control for during speech (it seems impossible to preselect certain words while speaking, based on their syllabic complexity), L2 training focused on the production of consonant clusters and complex syllable structures specifically may lead to improvements in overall L2 rhythm production. As phonetic training in the L2 classroom tends to be restricted to the pronunciation of individual phonemes, additional training might be included of these two slightly larger phonetic units. In addition to our own findings, it is worth noting that previous work also suggested that specific prosodic training, although generally highly uncommon in formal L2 education, can improve L2 prosody production (e.g., Moyer, 1999; Ramírez Verdugo, 2006).

(3) A FOREIGN ACCENT DOES NOT AUTOMATICALLY OBSTRUCT COMMUNICATION

The perception studies reported in this dissertation showed that deviance in L2 prosody contributed in different ways to L1 perceptions of accentedness, comprehensibility, and intelligibility. Consequently, L2 teachers may find it relevant to distinguish between subjective and objective measures in class as well. L2 learners might be heartened by the idea that while completely native-like productions are often unattainable for the average adult L2 learner, communication between L2 and L1 speakers does not have to be obstructed as a result of reduced intelligibility, as shown in our study on the effect of deviance in pitch accent distributions on accentedness, comprehensibility, and intelligibility.

6.5 Limitations and future work

As specific limitations and suggestions for future work are discussed in detail in the individual empirical chapters of this dissertation, we will limit ourselves here to those aspects that transcend individual studies. One of the limitations of this dissertation that concerns both Chapters 2 and 4, is the fact that as DLS and SLD who speak the standard variety of both Spanish and Dutch were difficult to find, especially at the higher proficiency levels, and hence other, sometimes non-linguistic, factors could not be controlled for in our studies. While we ensured that learners were comparable based on their overall L2 proficiency, their knowledge of additional L2s, and any physical issues that might prevent them from participating optimally in our tasks, it was impossible to also control for factors such as musical ability, motivation, and language aptitude. While the combination of Dutch and Spanish was felicitous in this context because it is an under-investigated L1-L2 combination and the typological prosodic differences between both languages make cross-linguistic comparisons relevant, future work might use languages that have more L1 and L2 speakers, because that would facilitate the finding of such speakers and subsequently enable researchers to control for more additional speaker characteristics. In addition, the stature of these language may be taken into account as Dutch and Spanish undoubtedly differ in this respect, the latter being a language that is currently spoken by approximately 477 million L1 speakers (Instituto Cervantes, 2017), while the former has roughly 24 million L1 speakers (Nederlandse Taalunie, Meertens Instituut & Universiteit Gent, 2017).

Another relevant issue is the naturalness of speech after it has been subjected to transplantation and resynthesis techniques. As we aimed to keep our stimuli as natural as possible in the perception study reported in Chapter 3, we chose to avoid such techniques there, and controlled for other types of deviance by manipulating the correspondence between the elicitation and focus conditions. However, such a method is not always available, and especially difficult to apply when multiple prosodic cues are to be manipulated at once, as was the case in Chapter 5. Although time consuming, future work might attempt to use more natural speech stimuli, while simultaneously manipulating several prosodic cues.

6.6 Conclusion

As described in Chapter 1, the aim of this dissertation has been two-fold: in half of the empirical chapters (i.e., Chapters 2 and 4), we studied the acquisition of L2 prosody in Dutch and Spanish by comparing L2 leaners of both languages with L1 speakers. In addition to determining *how* L2 learners deviated from the L1 norm, we also investigated the effect of several factors that might explain *why* L2 learners did not produce target L2 prosody. Our studies showed that overall L2 proficiency level, learning direction, and language-specific properties such as syllable complexity all affect the degree of success with which learners produced the prosodic patterns of their L2. In the other half of the empirical chapters (i.e., Chapters 3 and 5), we examined how the prosodic errors made by Spanish L2 learners affected Dutch natives' perceptions of accentedness, comprehensibility, nativeness, and intelligibility. We concluded that different types of prosodic errors (e.g., in intonation, rhythm, and/or speech rate) contributed to measures of perception in different ways: While L1 listeners clearly rated L2 speech as more accented and more difficult to understand as a result of its prosodic deviance, intelligibility, and thus communication, was not in jeopardy.

Notes

[1] Technically, the data we collected for our L2 intonation production study (Chapter 2) also allow for a comparison of both learning directions (i.e., DLS vs. SLD). However, a reanalysis of the data is needed to enable corroboration of the Markedness Differential Hypothesis, as the results provide sufficient information to determine that the utterances by *both groups* are characterized by transfer effects that diminish with increasing proficiency, but in their current form they do not provide sufficient information to determine *which group* is more on target in their production of pitch accent distributions to mark focus.

References

Allen, G. D., & Hawkins, S. (1978). The development of phonological rhythm. In A. Bell, & J. Hooper (Eds.), *Syllables and segments* (pp. 172–185). Amsterdam, The Netherlands: North-Holland Publishing.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, *42*, 529–555.

Archibald, J. (1994). A formal model of learning L2 prosodic phonology. *Second Language Research*, *10*, 215–240.

Best, C. T. (1995). A direct-realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 171–204). Timonium, Maryland, United States: York Press.

Bunta, F., & Ingram, D. (2007). The acquisition of speech rhythm by bilingual Spanish- and English-speaking 4- and 5-year-old children. *Journal of Speech, Language, and Hearing Research*, *50*(4), 999–1014.

De Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, *10*(1), 7-21.

De Ruiter, L. E. (2015). Information status marking in spontaneous vs. read speech in story-telling tasks: Evidence from intonation analysis using GToBI. *Journal of Phonetics*, *48*, 29–44.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in Second Language Acquisition*, *19*, 1–16.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning 27*(2), 315–330.

Edmunds, P. (2009). *ESL speakers' production of English lexical stress: The effect of variation in acoustic correlates on perceived intelligibility and nativeness* (Doctoral dissertation). The University of New Mexico, New Mexico, United States.

Face, T. L. (2000). Prosodic manifestations of focus in Spanish. *Southwest Journal of Linguistics*, *19*, 45–62.

Flege, J. E. (1984). The detection of French accent by American listeners. *The Journal of the Acoustical Society of America*, *76*(3), 692-707.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, Maryland, United States: York Press.

Grabe, E., Watson, I., & Post, B. (1999). The acquisition of rhythmic patterns in English and French. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 1201–1204). San Francisco, California, United States: University of California.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*, 201–223.

Howell, P., & Kadi-Hanifi, K. (1991). Comparison of prosodic properties between read and spontaneous speech material. *Speech Communication, 10*, 163–169.

Hualde, J. I. (2005). *The sounds of Spanish*. Cambridge, Massachusetts, United States: Cambridge University Press.

Instituto Cervantes. (2017). *El español: una lengua viva* ('Spanish: a language that is alive', in Spanish). Retrieved on 20 October 2017 from: https://cvc.cervantes.es/lengua/espanol_legua_viva/pdf/espanol_ lengua_viva_2017.pdf

Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, *58*(2), 541-573.

Krahmer, E., & Swerts, M. (2001). On the alleged existence of contrastive accents. *Speech Communication*, *34*, 391–405.

Laan, G. (1997). The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a read speaking style. *Speech Communication, 22*, 43–65.

Li, A., & Post, B. (2014). L2 acquisition of prosodic properties of speech rhythm. *Studies in Second Language Acquisition*, *36*(2), 223-255.

Magen, H. S. (1998). The perception of foreign-accented speech. *Journal of phonetics*, *26*, 381–400.

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*, *32*(4), 543-563.

Mennen, I. (2015). Beyond segments: Towards a L2 intonation learning theory. In E. Delais-Roussarie, M. Avanzi, & S. Herment. *Prosody and language in contact* (pp. 171-188). Heidelberg, Germany: Springer-Verlag.

Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition*, *21*(1), 81-108.

Munro, M. J. (1995). Nonsegmental factors in foreign accent: Ratings of filtered speech. *Studies in Second Language Acquisition 17*(1), 17-34.

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, *38*, 289–306.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *49*, 285–310.

Munro, M. J., & Derwing, T. M. (2001). Modelling perceptions of the accentedness and comprehensibility of L2 speech the role of speaking rate. *Studies in Second Language Acquisition, 23*, 451–468.

Nava, E., & Zubizarreta, M. L. (2010). Deconstructing the nuclear stress algorithm: Evidence from second language speech. In N. Erteschik-Shir, & L. Rochman (Eds.) *The sound patterns of syntax* (pp. 291–316). Oxford, United Kingdom: Oxford University Press.

Nederlandse Taalunie, Meertens Instituut, & Universiteit Gent. (2017). *Staat van het Nederlands: Over de taalkeuzes van Nederlanders en Vlamingen in het dagelijks leven* ('The state of Dutch: On the language choices of the Dutch and Flemisch in daily life', in Dutch). Retrieved on 20 October 2017 from: http://taalunieversum.org/sites/tuv/files/downloads/Staat_van_het_Nederlands_-_Onderzoeksrapport_ 2017.pdf

Nooteboom, S. G., & Kruyt, J. G. (1987). Accents, focus distribution, and the perceived distribution of given and new information: An experiment. *Journal of the Acoustical Society of America*, *82*, 1512–1524.

Ordin, M., & Polyanskaya, L. (2015). Acquisition of speech rhythm in a second language by learners with rhythmically different native languages. *The Journal of the Acoustical Society of America*, *138*(2), 533-544.

Özçelik, Ö. (2016). The Prosodic Acquisition Path Hypothesis: Towards explaining variability in L2 acquisition of phonology. *Glossa: a Journal of General Linguistics*, *1*(1), 1-48.

Payne, E., Post, B., Prieto, P., Vanrell, M., & Astruc, L. (2012). Measuring child rhythm. *Language and Speech*, *55*, 203–229.

Polyanskaya, L., Ordin, M., & Busà, M. G. (2017). Relative salience of speech rhythm and speech rate on perceived foreign accent in a second language. *Language and Speech*, *60*(3), 333-355.

Prieto, P., Vanrell, M., Astruc, L., Payne, E., & Post, B. (2012). Phonotactic and phrasal properties of speech rhythm: Evidence from Catalan, English, and Spanish. *Speech Communication*, *54*, 681–702.

Ramírez Verdugo, M. D. (2006). A study of intonation awareness and learning in non-native speakers of English. *Language Awareness*, *15*(3), 141-159.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *The Journal of the Acoustical Society of America*, *105*(1), 512-521.

Rasier, L., & Hiligsmann, P. (2009). Exploring the L1-L2 relationship in the L2 acquisition of prosody. *Online proceedings of first and second languages: Exploring the relationship in pedagogy-related contexts*, Oxford. Retrieved from: http://www.education.ox.ca.uk/wordpress/wp-content/uploads/2010/08/Rasier Hiligsmann.doc

Reichle, R., & Birdsong, D. (2014). Processing focus structure in L1 and L2 French: L2 proficiency effects on ERPs. *Studies in Second Language Acquisition*, *36*, 535–564.

Saito, K., Trofimovich, P, & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, *37*(2), 217-240.

Schairer, K. E. (1992). Native speaker reaction to non-native speech. *The Modern Language Journal*, *76*, 309–319.

Schmidt, E., & Post, B. (2015). The development of prosodic features and their contribution to rhythm production in simultaneous bilinguals. *Language and Speech*, *58*(1), 24-47.

Sluijter, A. M. C. (1995). *Phonetic correlates of stress and accent* (Doctoral dissertation). Leiden University, Leiden, The Netherlands.

Swerts, M., Krahmer, E., & Avesani, C. (2002). Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics*, *30*, 629–654.

Swerts, M., Strangert, E., & Heldner, M. (1996). F0 declination in read-aloud and spontaneous speech. In *Proceedings of the Fourth International Conference on Spoken Language*, *96*, 1501–1504.

Swerts, M., & Vroomen, J. (2015). Accent shifts in spoken noun phrases affect verification latencies of listeners in Dutch but not Canadian French. *Journal of Phonetics*, *52*, 170–182.

Swerts, M., & Zerbian, S. (2010). Intonational differences between L1 and L2 English in South Africa. *Phonetica*, *67*, 127–146.

Van Els, T., & De Bot, K. (1987). The role of intonation in foreign accent. *The Modern Language Journal*, *71*, 147–155.

Van Heuven, V. J. J. P., Kruyt, J. G., & De Vries, J. W. (1981). Buitenlandsheid en begrijpelijkheid in het Nederlands van buitenlandse arbeiders, een verkennende studie (Nonnativeness and intelligibility of foreign workers in Dutch, an explorative study, in Dutch). *Forum der Letteren, 22*, 171–178.

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, *35*, 501–522.

Publication List

Publication list

Journal publications (peer-reviewed)

Van Maastricht, L., Krahmer, E., & Swerts, M. (2016a). Prominence patterns in a second language: Intonational transfer from Dutch to Spanish and vice versa. *Language Learning, 66*(1), 124-158.

Van Maastricht, L., Krahmer, E., & Swerts, M. (2016b). Native speaker perceptions of (non-)native prominence patterns: Effects of deviance in pitch accent distributions on accentedness, comprehensibility, intelligibility, and nativeness. *Speech Communication*, *83*, 21-33.

Van Maastricht, L., Krahmer, E., Swerts, M., & Prieto, P. (2018). Learning direction matters: A study on L2 rhythm acquisition by Dutch learners of Spanish and Spanish learners of Dutch. Accepted for publication in *Studies in Second Language Acquisition.*

Van Maastricht, L., Zee, T., Krahmer, E., & Swerts, M. (2017). The interplay of prosodic cues in a second language: How the speech rate, intonation, and rhythm of speech by Spanish learners of Dutch contributes to L1 Dutch perceptions of accentedness and comprehensibility. Under revision at *Language and Speech.*

Papers in conference proceedings (peer-reviewed)

Van Maastricht, L., Krahmer, E., & Swerts, M. (2015). Perceptual effects of deviance in pitch accent distributions in L1 and L2 Dutch. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, United Kingdom: The University of Glasgow.

Van Maastricht, L., Krahmer, E., Swerts, M., & Prieto, P. (2016). Learning L2 Rhythm: does the direction of acquisition matter? In *Proceedings of Speech Prosody 2016* (pp. 974-978). Boston, United States: Boston University.

Van Maastricht, L., Zee, T., Krahmer, E., & Swerts, M. (2017). L1 perceptions of L2 prosody: The interplay between intonation, rhythm, and speech rate and their contribution to accentedness and comprehensibility. In *Proceedings of Interspeech 2017* (pp. 364-368). Stockholm, Sweden: Stockholm University.

Abstracts in conference proceedings (peer-reviewed)

Van Maastricht, L., Krahmer, E., & Swerts, M. (2013). Acquiring native-like intonation in Dutch and Spanish: Comparing the L1 and L2 of native speakers and second language learners. In *Proceedings of Phonetics, Phonology and Language Contact (PPLC) 2013*. Paris, France.

Van Maastricht, L., Krahmer, E., & Swerts, M. (2014). "Would you like red or white wine? - I would like red wine": Native speaker perception of (non-)native intonation patterns. In *Proceedings of Tone and Intonation in Europe (TIE) 2014*. Utrecht, The Netherlands: Utrecht University.

Osório Côrtes, P., **Van Maastricht, L.,** Krahmer, E., & Swerts, M. (2015). Focus marking by means of pitch accents in Brazilian Portuguese. In *Proceedings of Architectures and Mechanisms for Language Processing (AMLaP) 2015* (p. 83). Valletta, Malta: University of Malta.

Van Maastricht, L., Krahmer, E., & Swerts, M. (2016). L2 Rhythm Acquisition: the question of learning direction. In *Proceedings of New Sounds 2016*. Aarhus, Denmark: Aarhus University.

Van Maastricht, L., Krahmer, E., Swerts, M., & Prieto, P. (2017). L2 Rhythm Production & Perception in Spanish and Dutch: The interplay with learning direction, fluency and intonation. In *Proceedings of 2nd workshop on Second Language Prosody*. Graz, Austria: University of Graz.

Acknowledgements

Acknowledgments

Thank you…*

Marc Swerts & Emiel Krahmer for your generous and inspirational advice and support. I have learned so much thanks to your supervision, not only about doing research, but also about the kind of supervisor I would like to be myself.

Ad Backus, Aoju Chen, Núria Esteve-Gibert, Sónia Frota, & Ineke Mennen for being in my dissertation committee, taking the time to carefully read my thesis, and being so constructive in your comments.

Martijn Goudbeek, Maria Mos, & Tim Zee for being willing to read part of my thesis and asking me tough questions during my try-out defense.

Pilar Prieto for welcoming me so warmly during my research visit in Barcelona and for our pleasant and fruitful collaboration.

Santiago González Fuente, Iris Huebscher, Alfonso Igualada Pérez, Olga Kushch, and all my other temporary colleagues at the Grup d'Estudis de Prosòdia for making me feel at home in Barcelona and for all the great times we've had during my stay and since then!

Karin van Nispen & Marieke Hoetjes for being the best paranymphs a girl can have and actually making the preparation for my defense fun!

Hans Westerbeek for making my years in Tilburg a lot more entertaining (minions!) and interesting (thanks for all the spot-on music and series suggestions) and for making my journal articles and this dissertation so much nicer to look at!

My former & present colleagues at Tilburg University and Radboud Universiteit Nijmegen: you make work fun!

Elena Kireva & Paolo Mairano for their invaluable help in using Correlatore.

Everyone at the Escuela Oficial de Idiomas in Madrid and Barcelona, Fuentes Academia de Español, Radboud Universiteit Nijmegen, Rijksuniversiteit Groningen, Tilburg University, & la Universidad Complutense de Madrid who participated in my studies: without you this thesis would not exist and I appreciate your time and effort immensely.

* Of course, it is impossible to rank these acknowledgements according to either my gratitude or the recipients' contribution. Therefore, being the utterly structured person that I am, I have decided to proceed in reversed chronological order, bar one exception.

Johanna Sattler, Arthur Verbiest, & Astrid van Winden for your help in organizing data collection at the Escuela Oficial de Idiomas in Madrid and Barcelona and welcoming me there so warmly.

María José Calvo González for helping me to organize data collection at the Universidad Complutense de Madrid.

Lex Bouterse for your solid advice before the start of my PhD project.

Juan Alba Durán, Teresa Alcacio de Nordentoft, Reineke Bok-Bennema, Azahara Cuesta García, Michelle de Jong, Bob de Jonge, Brigitte Kampers-Manhe, Arie Molendijk, Gerdientje Oggel, Jaap van Os, & Audrey Rousse-Malpat for being such lovely colleagues at the Department of Romance Languages and Cultures at Rijksuniversiteit Groningen and inspiring me then (and during my studies) to give research a try despite my fears!

Lummy van Wermeskerken for being a true friend and seeing my capabilities before I could. I still have all the postcards you sent me over the years!

Denise van Dijk & Jorith Metsemakers for always allowing me to be nothing more and nothing less than myself. Being with you feels like coming home.

My family for your support and love for as long as I can remember.

Job & Aylin for dutifully participating in my experiments when I asked you to, but never asking me if the thesis was almost done yet (you have no idea how rare that is!).

Mama en papa for being there 24/7 and reminding me that "je niet meer kunt doen dan je best".

Robbert for always wanting to be on my team.

Summary

Summary

The melodic and rhythmic characteristics of a language play a substantial role in what this language overall sounds like. Yet, in foreign language (L2) classrooms these aspects of the L2 often receive little to no attention. Similarly, how foreign language (L2) learners acquire these essential aspects of a language has remained largely unexplored in research. In an attempt to address this knowledge gap, half of the studies reported in this dissertation concerned the production of L2 intonation and rhythm by Dutch learners of Spanish (DLS) and Spanish learners of Dutch (SLD). In those studies, we examined how the L2 intonation and rhythm of these two learner groups differed from the L1 norm, but also attempted to explain why this deviance occurred. In the other half of the studies reported in this dissertation, we explored whether errors in prosody production by L2 learners may affect L1 listeners' judgments about the L2 speaker, and whether they even lead to actual miscommunication.

Thus, in **Chapter 2** we reported the results of an experiment in which DLS and SLD produced utterances that were varied in focus structure in order to investigate whether both learner groups were able to produce native-like intonation in varying contexts and whether the extent to which they were successful was affected by their overall L2 proficiency. We found that both DLS and SLD transfer the intonation patterns that are typical of their L1 to their L2, but that as their proficiency level increases, the transfer effects decrease. Another aim of the study was to determine whether the intonational structure of the L2 of the participants could also influence the intonation of their L1, in other words, whether there is more than just one transfer direction. The results showed that as L2 learners became more proficient, their L2 indeed started to influence their L1.

Chapter 3 explored how deviance in the use of intonation to mark focus by L1 Dutch and SLD affects four different measures of L1 Dutch perceptions. The first experiment used unedited fragments collected for Chapter 2 of this dissertation produced by L1 speakers of Dutch, more proficient SLD, and less proficient SLD. Consequently, these speech samples did not only differ in their pitch accent distributions, but also in segmental accuracy, and possibly other suprasegmental properties. The results showed that L1 Dutch listeners considered speech increasingly more difficult to understand, more foreign-accented and less native-like as the proficiency level of the speaker decreased. In the second experiment, we asked L1 Dutch listeners to indicate which audio fragment sounded more natural within a specific focus context; the one in which the elicitation context was congruent with the focus condition or the one that was actually elicited in the opposite focus condition. The results showed that Dutch natives can distinguish between utterances based on pitch accent placement alone and can make the corresponding nativeness judgments. Finally, a reaction times task revealed that L1 Dutch listeners do not respond slower to L1 speech containing atypical pitch accent distributions that to speech containing typical pitch accent distributions. While they do respond

slower to L2 speech overall, and even more to less proficient L2 speech than to more proficient L2 speech, we attribute this delay in processing to segmental, rather than suprasegmental errors in the speech.

The study reported in **Chapter 4** examined whether DLS and SLD were equally capable of acquiring native-like speech rhythm, with the aim of determining whether *learning* direction (not to be confused with *transfer* direction), as well as syllable structure complexity, affects this process. We predicted that the speech rhythm of Dutch is more difficult to acquire for Spanish learners than vice versa. Our results showed that the target L2 rhythm was indeed easier to acquire for DLS than for SLD, suggesting that learning direction is an additional factor that plays a role in the L2 acquisition process. We also found an effect of syllable structure on the target-like production of L2 speech rhythm, in that L2 learners performed systematically more on-target when they produced sentences in which syllable structure was kept relatively simple than when they produced sentences in which relatively more complex syllables structures were used. This effect was more noticeable for SLD than for SLD. Therefore, we consider the L2 rhythm acquisition process to be an interplay between several phonetic properties, including phonotactic constraints and constraints on consonant cluster production.

In **Chapter 5**, we investigated the relative contribution of deviance in speech rate, intonation, and rhythm in speech by SLD to L1 Dutch perceptions of accentedness and comprehensibility. Speech resynthesis techniques were used to create the speech stimuli. The results showed that accentedness and comprehensibility were affected differently by deviance in (a combination of) intonation, speech rate, and rhythm: Accentedness was only affected by the individual manipulations of intonation and speech rate, whereas comprehensibility was affected by those two individual manipulations, but also by several interactions between those two, and even all three, prosodic cues. No main effect of rhythm was found on either accentedness or comprehensibility.

Taken together, the experiments reported in this dissertation show that there are several factors affecting successful L2 intonation and rhythm production, for instance, the direction in which learning occurs (i.e., from Dutch to Spanish or vice versa), the syllable complexity constraints of the target language, the nature of the prosodic cue, and the proficiency level of the L2 learner. With respect to L2 prosody perception, this dissertation shows that L1 perceptions vary depending on the prosodic cue in question. In addition, it is found that different perception measures are unequally affected by atypicality in a specific prosodic feature. Hence, future work on L2 prosody production might focus on how the factors influencing L2 prosody acquisition interact, while where L2 prosody perception is concerned, studies in which different perception measures are investigated in combination with multiple prosodic features will yield conclusions with higher external validity. From a more practical standpoint, we feel our studies have yielded three suggestions that are useful for L2 learners, as well as their teachers: 1) Learning the intonation and rhythm of a language is difficult, but not impossible as it will co-occur naturally with increasing general L2 proficiency, 2) Training may improve L2 prosody skills, so there are specific steps that can be taken to develop L2 intonation and rhythm, and 3) While atypical intonation and rhythm may lead to a foreign accent in the L2, this does not necessarily impede actual communication with L1 speakers.

TiCC PhD Series

TiCC PhD Series

1. **Pashiera Barkhuysen.** *Audiovisual Prosody in Interaction.* Promotors: M.G.J. Swerts, E.J. Krahmer. Tilburg, 3 October 2008.

2. **Ben Torben-Nielsen.** *Dendritic Morphology: Function Shapes Structure.* Promotors: H.J. van den Herik, E.O. Postma. Co-promotor: K.P. Tuyls. Tilburg, 3 December 2008.

3. **Hans Stol.** *A Framework for Evidence-based Policy Making Using IT.* Promotor: H.J. van den Herik. Tilburg, 21 January 2009.

4. **Jeroen Geertzen.** *Dialogue Act Recognition and Prediction.* Promotor: H. Bunt. Co-promotor: J.M.B. Terken. Tilburg, 11 February 2009.

5. **Sander Canisius.** *Structured Prediction for Natural Language Processing.* Promotors: A.P.J. van den Bosch, W. Daelemans. Tilburg, 13 February 2009.

6. **Fritz Reul.** *New Architectures in Computer Chess.* Promotor: H.J. van den Herik. Co-promotor: J.W.H.M. Uiterwijk. Tilburg, 17 June 2009.

7. **Laurens van der Maaten.** *Feature Extraction from Visual Data.* Promotors: E.O. Postma, H.J. van den Herik. Co-promotor: A.G. Lange. Tilburg, 23 June 2009 (cum laude).

8. **Stephan Raaijmakers.** *Multinomial Language Learning.* Promotors: W. Daelemans, A.P.J. van den Bosch. Tilburg, 1 December 2009.

9. **Igor Berezhnoy.** *Digital Analysis of Paintings.* Promotors: E.O. Postma, H.J. van den Herik. Tilburg, 7 December 2009.

10. **Toine Bogers.** *Recommender Systems for Social Bookmarking.* Promotor: A.P.J. van den Bosch. Tilburg, 8 December 2009.

11. **Sander Bakkes.** *Rapid Adaptation of Video Game AI.* Promotor: H.J. van den Herik. Co-promotor: P. Spronck. Tilburg, 3 March 2010.

12. **Maria Mos.** *Complex Lexical Items.* Promotor: A.P.J. van den Bosch. Co-promotors: A. Vermeer, A. Backus. Tilburg, 12 May 2010 (in collaboration with the Department of Language and Culture Studies).

13. **Marieke van Erp.** *Accessing Natural History. Discoveries in data cleaning, structuring, and retrieval.* Promotor: A.P.J. van den Bosch. Co-promotor: P.K. Lendvai. Tilburg, 30 June 2010.

14. **Edwin Commandeur.** *Implicit Causality and Implicit Consequentiality in Language Comprehension.* Promotors: L.G.M. Noordman, W. Vonk. Co-promotor: R. Cozijn. Tilburg, 30 June 2010.

15. **Bart Bogaert.** *Cloud Content Contention.* Promotors: H.J. van den Herik, E.O. Postma. Tilburg, 30 March 2011.

16. **Xiaoyu Mao.** *Airport under Control.* Promotors: H.J. van den Herik, E.O. Postma. Co-promotors: N. Roos, A. Salden. Tilburg, 25 May 2011.

17. **Olga Petukhova.** *Multidimensional Dialogue Modelling.* Promotor: H. Bunt. Tilburg, 1 September 2011.

18. **Lisette Mol.** *Language in the Hands.* Promotors: E.J. Krahmer, A.A. Maes, M.G.J. Swerts. Tilburg, 7 November 2011 (cum laude).

19. **Herman Stehouwer.** *Statistical Language Models for Alternative Sequence Selection.* Promotors: A.P.J. van den Bosch, H.J. van den Herik. Co-promotor: M.M. van Zaanen. Tilburg, 7 December 2011.

20. **Terry Kakeeto-Aelen.** *Relationship Marketing for SMEs in Uganda.* Promotors: J. Chr. van Dalen, H.J. van den Herik. Co-promotor: B.A. Van de Walle. Tilburg, 1 February 2012.

21. **Suleman Shahid.** *Fun & Face: Exploring non-verbal expressions of emotion during playful interactions.* Promotors: E.J. Krahmer, M.G.J. Swerts. Tilburg, 25 May 2012.

22. **Thijs Vis.** *Intelligence, Politie en Veiligheidsdienst: Verenigbare Grootheden?* Promotors: T.A. de Roos, H.J. van den Herik, A.C.M. Spapens. Tilburg, 6 June 2012 (in collaboration with the Tilburg School of Law).

23. **Nancy Pascall.** *Engendering Technology Empowering Women.* Promotors: H.J. van den Herik, M. Diocaretz. Tilburg, 19 November 2012.

24. **Agus Gunawan.** *Information Access for SMEs in Indonesia.* Promotor: H.J. van den Herik. Co-promotors: M. Wahdan, B.A. Van de Walle. Tilburg, 19 December 2012.

25. **Giel van Lankveld.** *Quantifying Individual Player Differences.* Promotors: H.J. van den Herik, A.R. Arntz. Co-promotor: P. Spronck. Tilburg, 27 February 2013.

26. **Sander Wubben.** *Text-to-text Generation Using Monolingual Machine Translation.* Promotors: E.J. Krahmer, A.P.J. van den Bosch, H. Bunt. Tilburg, 5 June 2013.

27. **Jeroen Janssens.** *Outlier Selection and One-Class Classification.* Promotors: E.O. Postma, H.J. van den Herik. Tilburg, 11 June 2013.

28. **Martijn Balsters.** *Expression and Perception of Emotions: The Case of Depression, Sadness and Fear.* Promotors: E.J. Krahmer, M.G.J. Swerts, A.J.J.M. Vingerhoets. Tilburg, 25 June 2013.

29. **Lisanne van Weelden.** *Metaphor in Good Shape.* Promotor: A.A. Maes. Co-promotor: J. Schilperoord. Tilburg, 28 June 2013.

30. **Ruud Koolen.** *"Need I say More? On Overspecification in Definite Reference."* Promotors: E.J. Krahmer, M.G.J. Swerts. Tilburg, 20 September 2013.

31. **J. Douglas Mastin.** *Exploring Infant Engagement. Language Socialization and Vocabulary. Development: A Study of Rural and Urban Communities in Mozambique.* Promotor: A.A. Maes. Co-promotor: P.A. Vogt. Tilburg, 11 October 2013.

32. **Philip C. Jackson. Jr.** *Toward Human-Level Artificial Intelligence – Representation and Computation of Meaning in Natural Language.* Promotors: H.C. Bunt, W.P.M. Daelemans. Tilburg, 22 April 2014.

33. **Jorrig Vogels.** *Referential choices in language production: The Role of Accessibility.* Promotors: A.A. Maes, E.J. Krahmer. Tilburg, 23 April 2014.

34. **Peter de Kock.** *Anticipating Criminal Behaviour.* Promotors: H.J. van den Herik, J.C. Scholtes. Copromotor: P. Spronck. Tilburg, 10 September 2014.

35. **Constantijn Kaland.** *Prosodic marking of semantic contrasts: do speakers adapt to addressees?* Promotors: M.G.J. Swerts, E.J. Krahmer. Tilburg, 1 October 2014.

36. **Jasmina Marić.** *Web Communities, Immigration and Social Capital.* Promotor: H.J. van den Herik. Copromotors: R. Cozijn, M. Spotti. Tilburg, 18 November 2014.

37. **Pauline Meesters.** *Intelligent Blauw.* Promotors: H.J. van den Herik, T.A. de Roos. Tilburg, 1 December 2014.

38. **Mandy Visser.** *Better use your head. How people learn to signal emotions in social contexts.* Promotors: M.G.J. Swerts, E.J. Krahmer. Tilburg, 10 June 2015.

39. **Sterling Hutchinson.** *How symbolic and embodied representations work in concert.* Promotors: M.M. Louwerse, E.O. Postma. Tilburg, 30 June 2015.

40. **Marieke Hoetjes.** *Talking hands. Reference in speech, gesture and sign.* Promotors: E.J. Krahmer, M.G.J. Swerts. Tilburg, 7 October 2015

41. **Elisabeth Lubinga.** *Stop HIV. Start talking? The effects of rhetorical figures in health messages on conversations among South African adolescents.* Promotors: A.A. Maes, C.J.M. Jansen. Tilburg, 16 October 2015.

42. **Janet Bagorogoza.** *Knowledge Management and High Performance. The Uganda Financial Institutions Models for HPO.* Promotors: H.J. van den Herik, B. van der Walle, Tilburg, 24 November 2015.

43. **Hans Westerbeek.** *Visual realism: Exploring effects on memory, language production, comprehension, and preference.* Promotors: A.A. Maes, M.G.J. Swerts. Copromotor: M.A.A. van Amelsvoort. Tilburg, 10 February 2016.

44. **Matje van de Camp.** *A link to the Past: Constructing Historical Social Networks from Unstructured Data.* Promotors: A.P.J. van den Bosch, E.O. Postma. Tilburg, 2 Maart 2016.

45. **Annemarie Quispel.** *Data for all: How professionals and non-professionals in design use and evaluate information visualizations.* Promotor: A.A. Maes. Copromotor: J. Schilperoord. Tilburg, 15 June 2016.

46. **Rick Tillman**. *Language Matters: The Influence of Language and Language Use on Cognition.* Promotors: M.M. Louwerse, E.O. Postma. Tilburg, 30 June 2016.

47. **Ruud Mattheij.** *The Eyes Have It.* Promoteres: E.O. Postma, H. J. Van den Herik, and P.H.M. Spronck. Tilburg, 5 October 2016.

48. **Marten Pijl.** *Tracking of human motion over time.* Promotors: E. H. L. Aarts, M. M. Louwerse. Copromotor: J. H. M. Korst. Tilburg, 14 December 2016.

49. **Yevgen Matusevych.** *Learning constructions from bilingual exposure: Computational studies of argument structure acquisition.* Promotor: A.M. Backus. Co-promotor: A.Alishahi. Tilburg, 19 December 2016.

50. **Karin van Nispen.** *What can people with aphasia communicate with their hands?A study of representation techniques in pantomime and co-speech gesture.* Promotor: E.J. Krahmer. Co-promotor: M. van de Sandt-Koenderman. Tilburg, 19 December 2016.

51. **Adriana Baltaretu.** *Speaking of landmarks. How visual information inuences reference in spatial domains.* Promotors: A.A. Maes and E.J. Krahmer. Tilburg, 22 December 2016.

52. **Mohamed Abbadi.** *Casanova 2, a domain specific language for general game development.* Promotors: A.A. Maes, P.H.M. Spronck and A. Cortesi. Co-promotor: G. Maggiore. Tilburg, 10 March 2017.

53. **Shoshannah Tekofsky.** *You Are Who You Play You Are. Modelling Player Traits from Video Game Behavior.* Promotors: E.O. Postma and P.H.M. Spronck. Tilburg, 19 June 2017.

54. **Adel Alhuraibi.** *From IT-BusinessStrategic Alignment to Performance: A Moderated Mediation Model of Social Innovation, and Enterprise Governance of IT.* Promotors: H.J. van den Herik and Prof. dr. B.A. Van de Walle. Co-promotor: Dr. S. Ankolekar. Tilburg, 26 September 2017.

55. **Wilma Latuny.** *The Power of Facial Expressions.* Promotors: E.O. Postma and H.J. van den Herik. Tilburg, 29 September 2017.

56. **Sylvia Huwaë.** *Different Cultures, Different Selves? Suppression of Emotions and Reactions to Transgressions across Cultures.* Promotors: E.J. Krahmer and J. Schaafsma. Tilburg, 11 October, 2017.

57. **Mariana Serras Pereira.** *A Multimodal Approach to Children's Deceptive Behavior.* Promotor: M. Swerts. Co-promotor: S. Shahid Tilburg, 10 January, 2018.

58. **Emmelyn Croes.** *Meeting Face-to-Face Online: The Effects of Video-Mediated Communication on Relationship Formation.* Promotors: E.J. Krahmer and M. Antheunis. Co-promotor A.P. Schouten. Tilburg, 28 March 2018.

59. **Lieke van Maastricht.** *Second Language Prosody: Intonation and Rhythm in Production and Perception.* Promotors: E.J. Krahmer and M. Swerts. Tilburg, 9 May 2018.

ISBN 978-94-6295-903-3